American Statistical Association

Organized November 27, 1839
Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and other persons who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association. Information about the Association and membership application forms may be secured from the Secretary, R. L. Funkhouser, American University, Washington, D. C.

The 104th Annual Meeting of the American Statistical Association will be held in Cleveland, Ohio, December 28 to 30, 1942. With a few exceptions the sessions are open to the public.

Chapters of the Association have been organized in important cities in the United States. Information concerning the meetings of these chapters may be secured from the District Representatives listed on another page.

The Editors welcome the submission of articles and notes for possible publication in the JOURNAL. Manuscripts should be sent to the Editor, Journal of the American Statistical Association, American University, Washington, D. C. Books for review should be sent to the Review Editor, Glenn E. McLaughlin, American University, Washington, D. C. Authors who wish suggestions about the preparation of manuscripts and charts or information about editorial policies should address their inquiries to the Editor.

Subscription rate, $5.00 per annum. Single numbers of volumes 23 to 30 $1.50 per copy postpaid, except Proceedings Supplements to volumes 23 to 30, $2.00 per copy postpaid. Prices for sets and discounts for orders in quantity available on request.

Published Quarterly by the AMERICAN STATISTICAL ASSOCIATION
[illegible] American University, [illegible] in the Act of [illegible] 1925. Entered as [illegible]

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

MARCH, 1942 · $1.50 PER COPY · $6.00 PER ANNUM · VOL. 37 · NO. 217

CONTENTS

ARTICLES

PROCEEDINGS OF THE 103rd ANNUAL MEETING

BOOK REVIEWS

Continued on next page

* For TNEC Monographs Nos. 5 and 14 see Brown and Perlman.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

JUNE, 1942 · $1.50 PER COPY · $6.00 PER ANNUM · VOL. 37 · NO. 218

CONTENTS

ARTICLES

BOOK REVIEWS

Continued on next page

BOOK REVIEWS (*Continued*)

* For TNEC Monograph No 39 see Cook.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

SEPTEMBER, 1942 · $1.50 PER COPY · $6.00 PER ANNUM · VOL. 37 · NO. 219

CONTENTS

ARTICLES

BOOK REVIEWS

Continued on next page

BOOK REVIEWS (*Continued*)

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

DECEMBER, 1942 · $1.50 PER COPY · $6.00 PER ANNUM · VOL. 37 · NO. 220

CONTENTS

ARTICLES

BOOK REVIEWS

Continued on next page

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 217 MARCH, 1942 *Volume 37*

GOVERNMENT AND THE STATISTICIAN*

BY WINFIELD W. RIEFLER

In 1917 the statistician invaded Washington. He came to man the war agencies, to build up a service of economic reconnaissance for the more effective mobilization of our resources. He found there little data with which to work and few colleagues with whom to consult. Those colleagues that he found were centered largely in the Census Bureau. Though small in number, they were an able group, pioneers in the long endeavor to construct a more exact description of the society in which we live. Frequently the volunteer who came to Washington in 1917 did not know that he was a statistician and did not become aware of that fact for some time afterward. He was, usually, an economist who had had little opportunity to familiarize himself with factual data in quantitative form. He had, however, a great faith in facts and soon acquired an enthusiasm for the questionnaire. When he departed, he had been exposed to a new range of problems and to a factual technique for dealing with those problems. He left behind him a mass of unassimilated data for other statisticians to digest.

During the past eighteen months, a second army of statisticians has invaded Washington. They, too, have come to man the emergency agencies, to provide the fact finding that is the basis for industrial mobilization in a modern war. They did not come to a statistical wilderness, however, nor have they lacked colleagues in number to welcome them. They arrived at a humming center of statistical work, rich in data, and widely staffed by specialized experts. Those who arrived, moreover, have differed from the earlier invaders. They have not been "rookies," but skilled workers already versed in statistical techniques.

The contrast in these two situations, separated by less than a generation, portrays a development of profound significance to statisticians, particularly economic statisticians. In the interval between the two wars, statistics have "come of age" and the statistician, as a species, has multiplied. The pattern of this phenomenon has not been peculiar

* Presidential address presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 29, 1941.

to this country. The same development can be traced, subject to local variations, over wide areas of the globe. In a very real sense, nevertheless, it has been an American phenomenon. In the area of economic statistics, faith in the statistical approach, energy in the compilation of new data, and reliance on statistical findings were early in making their appearance in this country. Throughout the past twenty-four years, this faith, energy and reliance have tended progressively to spread outside of our borders. They have been contagious and have quickened the adoption of comparable procedures abroad. One may say that they have constituted a cultural export. Just as faith in the new industrial techniques, and knowledge of the machine on which they were based, constituted a leading cultural export of Great Britain during the nineteenth century, we have exported an enthusiasm for the statistical approach in the twentieth. As a cultural export, this development may be ranked with our earlier missionary activities in Africa and the Far East, and with the picture of American life spread throughout the world by the movies during more recent years.

GOVERNMENT AS A FACTOR IN THE EXPANSION OF STATISTICAL ACTIVITY

It is no accident that the statistician first came to Washington in numbers in 1917, for it was then that government in this country first embarked on its course of intimate contact with, and direct intervention in, our economic life, a course which has been pursued continuously during the interval. Fundamentally, the concurrent expansion of activity in the field of economic statistics has reflected this development. Throughout the period government has been the great manufacturer of statistics. This does not imply that statistical activity has been confined to the public services nor that the statistician had no contribution to make to our understanding of economic developments prior to 1917. His contribution might have been great during those earlier years had he had the opportunity. However, in that era of economic liberalism—of free economic activity with as little interference as possible by the state—his contributions were necessarily limited because there was no single organization sufficiently powerful, or in possession of sufficient funds, to finance the acquisition and compilation of data on which statistical work depends.

All this was changed in 1917. The first World War was the occasion for the acceptance on the part of the government of widespread responsibility for participation in the conduct of our internal economic affairs. This participation changed its form, but did not end, at the close of the war. It has continued, in fact, and grown during the inter-

vening years. It has had mixed origins, reflecting, on the one hand, the voluntary initiation of measures directed toward the improvement of social conditions, and, on the other, a response to new problems imposed by the growing tension of the times. In the post-war period the government was forced to assume a degree of responsibility for economic readjustment and reconstruction. After the great collapse, it was government that was called upon to create some sort of economic order out of the chaos of 1932. In the succeeding era of world-wide economic disintegration and increasingly aggressive economic nationalism, agencies of economic intervention have tended to multiply, at first in support of exposed elements in the economy, and more recently to mobilize the economy for total war.

Taking the twenty-four years as a whole, this great growth in governmental functions bulks larger than any other single factor in the expansion of statistical activities. The statistician has found in government both sources of new data and an audience intensely interested in such findings as he was able to make. Government has certainly not been the *sole* source of new statistical data during this period, nor has it enjoyed any monopoly in the application of statistical analyses to the elucidation and understanding of current problems. In a very important degree, in fact, pioneering in the development of economic, financial and social statistics, in statistical techniques, and in their application has continued to originate in non-public agencies. Witness, for example, the recent innovation of the so-called public opinion polls. It remains true nevertheless, that in almost all phases of their activities, statisticians have been heavily dependent on government support. Public funds have financed in great part the collection of statistical data; public bodies predominantly have fostered the development of new statistical series, and public agencies have relied more heavily than others on the use of statistical guides for the development of administrative policies.

There are several reasons why this should be so. The cost of collecting mass statistics is frequently prohibitive. It is only a public agency supported by public funds that is in a position to decide to collect such basic data as are found, for example, in the Census Bureau. Furthermore, much of the statistical data on which we have come to rely, originated not as a result of a conscious decision to undertake their compilation, but rather as a necessary by-product of public administration. Our import and export statistics are a case in point, constituting essentially a by-product of customs administration. Thirdly, the activities of public administrative agencies are usually comprehensive in scope and their records are usually open to public inspection. In

consequence, statisticians have found in these records rich sources of statistical data. Public agencies, finally, have tended to place administrative reliance on statistical findings in the development of public policies because the field of their responsibility is broad. They have found that they could not afford to rely on opinion or impression when more objective bases for decision were available. This has meant, in the turmoil of the last quarter century during which government has tended to intervene ever more actively in economic life, a rapid stimulus to the availability, compilation and use of statistical data, especially data of economic significance.

Although the trend toward public participation in economic life appeared later in this country than in most of the other major countries, and rarely in so extreme a form, it had as a by-product, almost from the first, a parallel development of statistical activity. Because of our huge size and of the diversity of our internal conditions, there was earlier appreciation here of the need on the part of public authorities for objective factual data to describe conditions covering the country as a whole. They could not afford to place the same degree of reliance on experience, intuition, and impression as the authorities of smaller or more homogeneous states. Because of our more decentralized form of government, with powers and activities widely distributed between the Federal government and the states, public records have always tended to be more voluminous here than abroad. Our statisticians enjoyed, in consequence, earlier access to a wide variety of existing data capable of throwing at least some light on new problems as they emerged. Underlying both of these developments has been the American temperament, i.e., the willingness of the American to furnish information. In this country we have not experienced the same resistance to the gathering of statistical data, particularly data voluntarily supplied, as in many countries abroad.

Given these favorable conditions and the opportunity to demonstrate his technique on a large scale, the American statistician "came through." He was a success. With scattered data, often fragmentary, at his disposal, he was able to block out the order of magnitude of important elements in our economy about which we had hitherto been ignorant. With more adequate data he was able, frequently, to furnish fairly precise guides to decision. In a world characterized by increasing uncertainty, he was able constantly to whittle away at the area of the unknown.

These activities varied widely both with respect to the agencies by which they were sponsored and the diverse sources of data which were subjected to exploitation. They fall, nevertheless, into a general chrono-

logical pattern, that reflects in its shifts of emphasis the changing focus of government with respect to economic problems. Taking the period as a whole, there were three such major shifts—in 1920, in 1933, and in 1940—each corresponding to concurrent changes in the orientation of public policy. In large part these changes were initiated by the statistician, and thus represented a flexible response on his part to changing social needs. To a certain extent they were involuntary, however, forced by conditions over which he had little control. The statistician is a technician. He must have data, the collection and compilation of which are costly. Under these circumstances his major fields of operation are bound to reflect to a considerable extent areas of interest for which support of his activities can be obtained.

It is not surprising that, beginning in 1920, efforts were concentrated on the collection of banking data and on general economic data suitable for the construction of economic time series, nor that the research divisions of our monetary agencies were prominent in this role. This was the period in which we still conceived of our economy—domestic as well as international—as an integrated whole, maintained in equilibrium through a flexible system of prices. We thought of the business cycle as the great menace to the health of that whole. Our social thinking, consequently, was oriented toward the development of public policies, chiefly monetary policies, adapted to mitigate or eliminate the business cycle. It was assumed that these policies would operate over the whole economy and that, so long as that economy was maintained in stable equilibrium, its constituent parts would be secure. For the formulation of public policies directed toward these purposes, banking data were essential, together with economic time series, and the general purpose index numbers, based frequently on fragmentary but symptomatic data, which were constructed in this period, were in accord with this diagnosis. Though limited in coverage and subject to wide margins of error, they were useful in the interpretation of general trends in the economy, particularly those that were cyclical in character.

The principal exception to this generalization is to be found in the fields of agricultural statistics and coal mining statistics. All during the 1920's it was apparent that there was a malaise in agriculture and coal that was more than cyclical. In these two fields, consequently, statisticians pushed their work far beyond the construction of index numbers. They developed, in addition, data on a comprehensive basis capable of aggregation into relevant categories adapted for the detailed study of these specific segments of our economy.

Beginning in 1933 this latter type of statistical activity tended to generalize. In the long descent from the collapse of 1929 to the depths

of 1932, attention shifted from preoccupation with the state of the economy as a whole to concern over certain of its constituent parts. Public policy was directed more actively toward specific support for those parts on the assumption that such rehabilitation was a necessary prerequisite to the restoration of health to the whole. Statistically this shift in emphasis was reflected in such activities, to enumerate only a few, as the development of relief statistics, statistics of employment and unemployment by industrial categories, urban mortgage statistics, the segregation of industrial statistics of all kinds into durable and non-durable categories, and in a widespread attempt to describe statistically the forces that play on the construction industry.

A third shift of emphasis has characterized the past eighteen months, adapted to the mobilization of our resources for war. For this purpose, highly individualized data, frequently plant by plant, are needed, in addition to time series and to statistics descriptive of the economy by major economic segments. The Bureau of Labor Statistics index number of wholesale prices, for example, despite all its recent refinement, does not furnish sufficient data for the establishment of price ceilings. Nor are our industrial data, in all their elaboration, adequate for the purposes of the OPM.

While government has played a major role during the past generation both in providing support for and promoting the use of statistical data, particularly economic data, statistical activity has in no sense been restricted to public bodies, nor has it been confined to the area of economic phenomena. The period has been characterized by intense activity throughout the various fields where the use of statistics as a tool of analysis is capable of making a contribution. Within the field of economic data, much of the pioneering work, as, for example, in the field of income statistics, was initiated outside the government. Public agencies have not, in general, been prominent in the development and refinement of statistical methodology, nor in the elaboration of techniques of statistical analysis. They have made contributions in this area, but in general they have borrowed their methods and adapted their techniques from academic sources. Private business and finance also have participated in the general development of statistical activities, both in the interpretation of existing series and in the compilation of new data, as, for example, in the case of statistics gathered by trade associations. Financial institutions offered a major outlet for statisticians during the twenties. Market research has provided comparable opportunities in more recent years.

Public support, however, has remained at the core of the movement. Each new activity has increased the demand for statisticians in the

public service. Each new adoption of statistical techniques outside of the government has led to augmented demands on public agencies, such as the Census Bureau, for additional information. These demands have reacted in turn on universities which have been called upon to develop more refined methods of analysis and to augment their teaching programs in order to supply an adequate number of skilled technicians. In a sense, therefore, the army of statisticians that came to join that small band of pioneers already functioning in Washington, was never truly evacuated. The current heavy influx to man the defense agencies constitutes merely the latest phase of a phenomenon that has continued with few interludes throughout the intervening generation.

STATISTICAL CONTRIBUTIONS TO ECONOMICS

In view of the enormous resources thus placed at the disposal of the economic statistician during and since the first World War, it may properly be asked: What has the statistician done with these opportunities, for economics, for the social sciences? What progress has been made in the endeavor to describe quantitatively the economic environment in which we function? To what extent have we been successful in establishing more precise quantitative relationships between functionally related elements of this environment? What has been done with these findings? To what extent have they improved our wisdom as well as our knowledge? Is control of the business cycle more nearly within our grasp?

So far as success in describing the economy in which we function is concerned, the answer is, I think, decisive. The statistician has seized the opportunity to advance immeasurably our knowledge of the structure of our economy. So great is the extent of this advance, in fact, that no one person, no matter how expert, is any longer in a position to appraise it. The economist of 1917 could, if he chose to specialize, be reasonably conversant with the sources of all of the major statistics that were available to describe our economy. He could also retain in his memory a great bulk of the more significant of these facts. Today, it is doubtful whether the entire staff of any one of our most highly organized and competently equipped agencies for statistical research is in a comparable position. In spite of all the work that has been done to digest and to assemble into major categories the wealth of statistical materials now available and to reduce their movements to index numbers, the research agency of today must specialize in one or more fields of data. It relies upon intimate contact with specialists in other agencies to make good its deficiencies.

In a very important degree, many of the statistical aids which we

now accept as common knowledge, were completely lacking only a short time ago. Take, for example, *statistics of national income*. The monthly figures, on which so much current reasoning is based, are a product only of the last few years. Even continuous annual estimates did not become available until the twenties. Except for imports and exports of gold, we knew nothing of the current movement of international financial transactions until very recently, while the whole development of systematic data with respect to the balance of payments goes back also only to the early twenties. Before then, statistics of physical imports and exports comprised the bulk of our knowledge in that field. The statistician in 1917 did have a wholesale price index to follow, but an index much less revealing than that which we now scrutinize to measure the progress of war inflation. It is difficult to comprehend today, but nevertheless true, that we had no index of physical production whatever until the early 1920's. There were also little or no data on retail trade, on employment, or on wages. It is amusing, today, to re-read the most carefully compiled economic summaries of that earlier period from the point of view of the adequacy of the data that were used and their relevance to the inferences which they were cited to support. Bank clearings were considered a most important index of business activity, while railroad car loadings were studied for information on the state of retail trade.

In addition to these advances in sheer description of our economic environment, considerable progress has also been achieved in the establishment of functional relationships between diverse elements in the economy. The recent work in the important field of consumption statistics, for example, has provided a working basis for estimating the consumption habits of different classes in the community by income groups. We know in considerable detail the relative importance of different production groups. We now have records painstakingly traced back, breaking down our total economic effort into service production and commodity production, with further breakdowns of commodity production into nondurable consumers' goods, durable consumers' goods and producers' goods. We are able to assess the importance of maintenance activities in our economy, to measure the total of gross capital formation, and to establish the fact that a very large component in that total consists of replacement of existing facilities.

The relationships established by these basic statistical contributions have been highly suggestive. They have certainly laid the ground work for the new theories of employment, consumption, savings, and investment with which the economist is now working. There can be no question that they have added very greatly to our knowledge of the

functioning of our economy and of the mechanism of the business cycle. Opinion is less unanimous as to whether they have increased our wisdom with respect to its control.

OFFSETTING FACTORS

There is a different kind of question which may also be asked with respect to the influence exerted on the social sciences by the statistician and by the statistical approach. What has the statistician done to the social sciences? Has he contributed directly or indirectly, advertently or inadvertently to the social, economic and political disintegration that has culminated in this second World War?

The statistician, almost by definition, because of the material with which he works, cannot help but emphasize the immediate, the concrete, the specific, the tangible. Has he, by this very fact, in this period in which he has been successful, tended to direct attention from considerations and developments that have been long-run and intangible? Have the quantitative relations which he has demonstrated between different elements in the economy been so compelling in their power to attract attention that they have been over-emphasized as compared with other relationships, just as true and valid, but non-demonstrable statistically? In other words, has he made the economist more akin to an engineer and less responsible as a social philosopher?

The statistician must work with data, most of which have been national in their coverage. Predominantly he has developed these data in a form that would be useful to public agencies, the specific responsibilities of which were national. Has he, by this very fact, over-emphasized the national economy at the expense of the international organic whole of which it is but a part? Just how relevant are the concepts of *national* income, *national* production, in a world dedicated to non-totalitarian ideals? Have they not, in the absence of comparable world data of which they are essentially but a segment, tended to be over-emphasized, to become the central focus of analysis? Has the reiterated use of these statistical measures done something to the concepts behind them? Has it tended to make our economic thinking stop at the political border? In short, have we exported a flavor of our isolationism along with our enthusiasm for statistics?

There is, of course, no ultimate answer, pro or con, to probing questions such as these. We can only note them and resolve to overcome the faults implied in them as opportunity permits. With respect to one of them at least, however, we must plead guilty. There is no question but that we have over-concentrated on national data, that we tended to become complacent once our *internal* national economy had

been described. When one considers, on the one hand, the enormous efforts that have been devoted to the development of economic data during the past generation, and, on the other, the basic inter-relationships that knit the world's economy into a whole, it is difficult to explain the paucity of our statistical knowledge with respect to that whole. It can be accounted for in part, probably, by the fact that world data covering activities under the jurisdiction of many different political sovereignties are subject to wide gaps, and even when available, are frequently non-comparable in form. In our internal compilations, however, when we faced the same difficulties, we have frequently devised means to surmount them. Chiefly it is to be explained by the factor suggested above, namely, that so many of our research agencies have restricted their work to the development of series that would be useful to the public bodies for which they have worked, whose responsibilities were largely national. It is significant to note in this respect that we are indebted to the Economic and Financial Section of the League of Nations, a research agency responsible to an international body, for an important segment of such economic data as we *do* have on a world basis.

FUTURE LINES OF DEVELOPMENT

Adequate provision to make good this major gap in our information is one of the most important tasks facing the economic statistician today. Valuable as it is, the ground work already developed under the League and other international bodies barely scratches the surface of the problem. The world indexes now available, such as those on production, on international trade, and on stocks of primary materials, need to be broadened. We need economic time series comparable to those we have at home because the business cycle is not an internal phenomenon nor are its ravages confined to any one country. We need industry surveys on a world basis because the malaise which has stricken our basic industries—coal, textiles, cotton, wheat, coffee, and sugar—has in most cases a non-national origin. We need geographic area studies by industrial regions on a trans-national basis because these studies are essential to the rebuilding of our shattered world by the opening of markets and the provision of investment funds. Effort devoted to tasks such as these will pay well in dividends. It has been my privilege recently to go over some preliminary analyses of international trade statistics, integrated by regional geographic areas. I have rarely encountered a situation where a simple analysis, achieved mainly by the regrouping of familiar data, did more to undermine preconcep-

tions in understanding and to clarify intellectual areas that had hitherto been obscure.

The preliminary attack on this wide new field of statistical research is already under way. Under the impact of war, various agencies of the government have found the area of their responsibilities broadened to include activities outside our borders. They are being called upon to make decisions with respect to these activities and find themselves in need of non-national statistical data as a basis for these decisions. Mobilization of our industries for war has revealed deficiencies in our internal supplies that must be supplemented from outside. Economic warfare calls for intimate knowledge of critical elements in the economy of the enemy. Support of friendly powers requires information on which to base decisions with respect to priorities, to sustaining purchases, and to loans. For all of these purposes economic data are required, economic data that consist of more than a summary, nation by nation, of purely national statistics. For many purposes these data must be developed on an integrated basis, international or trans-national in character. These activities may go far before the war is over. It may be that in this war a basis will be laid for international statistical research comparable to the basis laid for national statistical research in 1917.

The trend, so established, will probably not be limited to the war period. If any forecast of the future is possible in this period of obscurity, it is that we have passed the crest of the era of national economic isolation. The post-war world, whatever its specific character, will be forced to recognize the problem of international economic interdependence, and to deal with it. That problem, furthermore, will be dealt with by public bodies in some form, such as international agencies, or joint commissions. We will have, therefore, all of the conditions requisite for the inauguration of successful research, the existence of a problem requiring the quantitative clarification which the statistician is able to provide, the presence of a fair body of data on which to work, the support of public agencies in possession of sufficient resources to prosecute effective inquiry and, finally, a pressing need on the part of those agencies for objective economic data on which to base their decisions. In the development of those data, essential for the preparation of plans and the formulation of policy, the statistician of the future, it is safe to predict, will have a unique opportunity to dedicate his energy to the reconstruction of an integrated world.

THE PROSPECT FOR 1942*

A NEW PRACTICE has been inaugurated of bringing to the membership a composite of the forecasts for the year 1942 of fifteen distinguished economic statisticians in business, government, and research. The results were presented and analyzed by W. Randolph Burgess, vice chairman of the National City Bank of New York and a former president of our Association.

The reporting statisticians agreed in predicting an increase in 1942 in industrial production, the individual forecasts ranging from an increase of one per cent by December 1942, to an increase of 15 per cent. The average estimate was for an increase of 9 per cent from December 1941, to December 1942, compared with an increase of 21 per cent during the year 1941. The estimates also agreed that both wholesale prices and the cost of living would be substantially higher a year from now. In the case of wholesale prices, estimates ranged from an increase of 13 to 29 per cent, with an average at 18 per cent. In the case of the cost of living, the range of estimate was from 6 to 17 per cent, with an average at 10 per cent.

Mr. Burgess said that these statisticians evidently expected the increase in war production to more than offset decreases in industries affected by priorities and shortages of materials. Actual entrance into the war may be expected to stimulate war production both in terms of the size of the program and the energy with which it is pursued.

In judging the effect of the war on prices from the point of view of the economics of the situation, inflationary forces are increased, as the amount of government spending expands. On the other hand, the still more important political factors in inflation are more favorable. The people will be more willing to tax themselves heavily and resist the efforts of pressure groups to raise agricultural prices and wages, and maintain non-defense spending. The political question, however, remains the dominating factor with respect to the country's ability to resist inflationary forces.

In World War I statisticians had to fight for a chance to demonstrate their usefulness. In the present war they are being eagerly sought after. Their responsibilities are correspondingly greater, and their success will depend more upon their power to find out what facts and interpretations of fact are needed by the people who make decisions than by their technical competence in refined statistical methodology.

* Abstract of material presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 30, 1940.

PRICE FIXING OF AGRICULTURAL PRODUCTS*

By Warren C. Waite
University of Minnesota

One of the major control problems of a war is that of restraining or preventing a disastrous rise in prices. At the outset it should be clearly recognized that price fixing in itself will not prevent a price rise unless it can be enforced by drastic means plus, of course, vigorous rationing. The mere fixing of prices does not provide a fundamental solution to the problem of rising prices. At best it is only an aid.

The fundamental cause of a general rise in prices is that the purchasing power in the hands of buyers rises relative to the quantities of goods available for purchase. The real solution to this problem is either to reduce the purchasing power of the buyers or to increase the volume of production of goods available for purchase. The latter, with the demands arising from a war effort, is clearly impossible in many lines of production. Agriculture, perhaps, constitutes a notable exception. The reduction of purchasing power is possible, and efforts in this direction may be expected to be carried out, but it is difficult in a complex economy to recapture all increases or to drive very deeply without creating grave injustices or hardships among classes in the community. Insofar as this balancing of goods and purchasing power is accomplished, the task of price fixing is made easier, and complete success in this respect would eliminate the need for the major portion of it. Even an over-all ceiling such as has been proposed by Mr. Baruch and others cannot prevent price rises if purchasing power is unrestrained or not taxed away. If you fix prices without changing the quantities of goods or purchasing power, then people will buy up everything at the established prices and have purchasing power left over. These free funds will surely lead to price rises.

Public opinion will do a great deal in helping to police a price-fixing scheme. In fact, we must depend pretty largely upon it to insure the success of any plan in the United States. There are, however, the difficulties that people don't know what prices ought to be, and there may still be among us those who are more interested in bettering their own situation than they are in the general welfare. Certain prices are not and cannot be very well known, for example, the price of agricultural land, of second-hand goods, or the services of domestic workers. These and others may easily be bid up. Lack of success here might easily

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 27, 1941.

Paper No. 1951, Scientific Journal Series, Minnesota Agricultural Experiment Station.

jeopardize the whole scheme. Assuming, however, a moderately successful control over purchasing power, it would appear that selective controls might be employed in the fixing of prices with considerable effectiveness. Among these selected prices will be certain agricultural products.

Any plan of price fixing should meet two general requirements. It must first be economically sound, that is there must be some real economic objective that can be brought about by the price fixing and that will be secured by it. Secondly, it must be administratively feasible. No plan, however desirable it may be from an economic viewpoint, is worth anything unless it can be administered. The difficulties of administration will be lessened to the extent to which the market can be made to function in its usual capacity during the period of price fixing. This suggests that whenever there is a choice of means for accomplishing a desired result, the one that provides a larger degree of automatic regulation through the agencies now available should be employed. The price administrator should conceive of his function as that of assisting the market to operate in a way beneficial to the common purpose, rather than one of supplanting or replacing the normal market.

In order to fix prices on agricultural products it is necessary to have in view the underlying economic purposes of such price fixing. Two important but partly contradictory purposes stand out at once. The first is to aid in the allocation of the costs of the war among the various groups in the community and in the case of agricultural price fixing the allocation between agriculture and the other groups. The second is to guide agricultural production during the war with a view of securing adequate supplies and adjusting agriculture to a position to be able to withstand the shock of returning to a peace-time economy. Preventing agricultural prices from rising too high will aid in restraining a general rise in all prices in several ways. Lower agricultural prices will lessen the cost of purchases of many goods bought by the Government. It will also keep the cost of food to workers lower, and this should lessen the pressure from the laboring classes for wage increases. Food is a large part of the worker's budget and a part in which he seems to be especially aware of price changes. Moreover, many persons will fail to secure appreciably higher incomes during the coming period, and for those persons a rise in food costs might create considerable hardships. Keeping agricultural prices down would also dampen the rise in agricultural incomes and lessen the increase in purchasing power in the hands of the farmer. This would curtail his ability to buy other

goods and in consequence his contribution toward inflation. There is thus in setting agricultural prices an income allocation between agriculture and the remainder of the community. This includes the tacit assumption that income or purchasing power in other hands is more diffused, less influential in producing price rises, or more subject to control than when in the hands of the farmers.

Price fixing also offers an opportunity to adjust agriculture toward an allocation of resources deemed more desirable and to accomplish this without the high cost to the remainder of society that might be involved if all prices were allowed to remain free to rise. Allocation of resources among productive enterprises in agriculture depends a great deal upon the relative prices of products. It is only the total volume of resources utilized by agriculture that depends in large part upon the level of prices. Total agricultural resources are not going to change very greatly either under high or low levels of agricultural prices, but considerable shifts among enterprises will take place with changes in the relative prices of products. Expansion of total agriculture production is likely to be nearly as great with a fair increase in income as it would be with an enormous increase in income.

To say that we want to adjust agriculture is not, however, to say that we are entirely certain just what that adjustment should be. The immediate problem is fairly clear, but the long-run problem is extremely uncertain. Immediately we need a schedule of requirements for the war effort together with the requirements for emergency stocks. There should naturally be a schedule in terms of absolute minimums and another of desirable quantities. We also need to recognize that it is more desirable to err on the side of too much rather than too little. We can always use up too much even if the process is painful for many, but deficits might be fatal.

This is not, however, to say that we should close our eyes to the future following the war even if our present view of this future is vague. It is obvious that we must currently produce what is essential to the prosecution of the war, but we must also keep in view the long-run as far as we are able. It is essential that we do this because we cannot avoid the difficulties of the future by the simple assumption that any increase in production induced now will be demanded by ourselves and our neighbors following the war. There may be needs from a nutritional viewpoint in our own population and from a humanitarian viewpoint for Europe, but whether a people in the course of a great adjustment to a peace economy and already burdened by the costs of a contribution to a war can freely give or will be willing to is an un-

decided question. The war, rather than solving our agricultural problem, will serve only to intensify it at the close, and present plans must keep an eye upon the possibilities of this future.

Any price fixed for agricultural products must face, perhaps in the law providing for price fixing and certainly among the farmers themselves, the handicap of comparison with "parity" prices. The idea or concept of parity is one which has come to be associated with the idea of right or justice in the minds of many agricultural leaders and the farmers themselves. Any ceiling set at parity levels would be likely to be accepted at once by the agricultural group as "fair," while it is conceivable that any price set below parity would be felt unjust if not actually objected to.

The idea of parity found its first legislative expression nearly a decade ago in the Agricultural Adjustment Act of 1933. In this Act it was declared to be the policy of Congress to reestablish prices to farmers at a level that would give agricultural commodities a purchasing power with respect to articles that farmers buy equivalent to the purchasing power of agricultural commodities in a base period. Briefly the essential calculation is to multiply the average price of the commodity in the base period, usually August 1909 to July 1914, by the index of the cost of things bought by the farmer, with the base of this latter index also from 1910 to 1914. This underlying idea has been expanded and developed in later agricultural legislation in various ways and has come to saturate agricultural proposals of many kinds.

Parity price and parity income have become an important ethical concept of what farmers ought to have in view of a situation in an earlier period. Price fixing in a war period should be unhampered by any legislative requirement to provide some historical relationship. The price administrators should be left free to operate as current conditions demand. The difficulty with parity prices as a criterion for ceilings in a war emergency is that they may be and probably are too high to provide the results selected as desirable or demanded in the fixing of many prices in such a period. In retrospect, the period 1910 to 1914 appears especially favorable for agricultural prices in relation to non-agricultural prices. To continue such a relationship in peacetimes is to invite the accumulation of forces tending to destroy this very relationship, and to force a set of price relationships of twenty-five years earlier on a price administrator is to handicap severely, if not to defeat completely, his efforts. Cost relationships among the commodities have changed greatly, and to provide parity prices for some commodities would perhaps stimulate their production greatly. To offset these shifts it would be necessary to establish price differences between prices

to enable the ceilings to exercise a directive function and a requirement of parity for the lowest prices in such a system might easily lift the entire level so high as to lose all possible inflationary control. If parity prices as minimum requirements in themselves promise to be faulty, then all *that can be said about the proposal to provide a* 110 per cent of parity minimum for any ceiling established is to make nearly certain the failure of agricultural price controls.

Assuming that the price administrator has been left with a fairly free hand, then in view of the purposes of agricultural price control which we have indicated the two principal problems for the administrator appear. The first is to establish a general level for agricultural prices relative to the remaining prices in the system, and the second is *to establish* the desired set of interrelations among the agricultural prices themselves. The criterion for the general level of agricultural prices is one of the amount of purchasing power which it is desirable to have in the hands of the farmers. While this cannot be determined precisely, two over-all measures will be useful. The first is the purchasing power *in the hands of the agricultural group as determined by* currently available data. This purchasing power should not be allowed to expand markedly nor should prices be set so low as to decrease it greatly. Here the administrator should probably tend to err on the upper side, as he probably will, because not to may easily be political suicide. *The second criterion is the proportion of consumers' income* being expended on foodstuffs. This sets an upper limit to be watched. The proportion of the expendable consumer purchasing power going for food should not be shrunk. Indeed it might properly be permitted to increase on the grounds that the great decreases in consumption must occur in *other* goods than in food. This would be an aid to the control of these other commodities. There are, however, many variations in the increase of purchasing power among the consumers and these increases must be brought under control primarily by other methods. In addition to these, of course, the relationship of agricultural income to consumer food expenditures is to be observed as an indication of what processors and retailers are doing. On the grounds of production response as a whole it does not appear probable that prices are likely *to be too low as judged by these criteria*. Patriotic motives will lead to nearly as large an increase in production as would occur if prices were left to skyrocket.

Once the general level of prices has been decided upon, there follows the problem of adjusting or fixing the individual prices about this level. The *purpose here* is primarily that of the adjustment of the production of the various agricultural commodities. Much of this could

be accomplished by simply making known to farmers the deficits, but this process can undoubtedly be facilitated by price manipulations. The criterion here is the rates at which these commodities are forthcoming relative to their requirements as determined by the overhead planning agency. Fortunately the data underlying all these decisions are pretty well at hand, and there is a vast quantity of administrative skill and judgment in the Department of Agriculture and among the farmers as a result of their own discussions of recent years.

The decision of the level at which the price shall be set having been made, the next problem becomes one of where and at what stage of the market it should be established, and what shall be done with respect to prices at other stages in the market or within the market area. One might establish only retail prices and allow farm prices to flow back from them. The retail price structure is, however, extremely complex and would probably be upset by even the simplest structure that could be devised. Another way would be to establish farm prices and allow the retail prices to be built up from them. Areal price structures for agricultural prices are also complex and offer difficulties of somewhat similar nature. The remaining major possibility is to set prices in a few of the major central markets and to have farm and retail prices flow from them. Central market wholesale prices have always been the great gauge by which agricultural products have moved and exchanged, and matters could be simplified if price fixing could proceed through them.

As long as the prices which are set in the central markets are reasonably well in line with prices that would have been established under ordinary competitive conditions in those markets, this procedure should work fairly well. If, however, prices are set that do vary widely with the price that would have prevailed without regulation, then marked changes may be brought about in the way and rates at which products are marketed, and this might force price fixing to be extended beyond the central market.

Dealing only with central market prices would permit at least for the time being retail prices to go unregulated. No retail regulation would be necessitated before retail prices began to rise relative to farm prices. Here again we would do well to fall back on the general principle of permitting the market to police itself as far as possible automatically. It seems probable that the consumers themselves would be able to exert a great corrective influence if only they are provided with proper information. Neglecting for the time being the perhaps crucial question of the ability of the consumer to judge quality in many products, much certainly can be accomplished by the provision of proper price

information. This would begin, of course, with lists of reasonable prices. A more important requirement would be that the merchant display the price which he paid for the goods sold along with his retail price to the consumer. If we accept the proposition that the consumer is entitled to know the truth about the goods which he buys, then the question is only whether the retailer's margin is his legitimate trade secret. Given this information, we could expect the consumer to do a pretty good job of protecting himself. A little shopping around would disclose those with suspicious margins or unusual costs of buying, and these could be subsequently called upon to verify their statements. Such a procedure would also work toward an improvement in marketing by giving an advantage to the more efficient and lower cost retailers.

The situation with respect to setting prices on the various grades of the commodity is essentially similar to that of areal designation of prices. As long as the level that is established for the dominant central market grade or grades is fairly close to the level that would otherwise have prevailed, the grade differentials should work out reasonably well. It is only when price rises have been greatly curtailed that grade differentials become necessary.

It is when we come to the field of production and the holding of commodities that price fixing may be expected to result in the greatest changes from the usual procedures of the market. To a large extent many of these changes will be unforseeable and unusual in their character. The decisions with respect to agricultural production are made by a multitude of farmers upon their individual farms. At certain times of the year, or perhaps in some cases less frequently, the farmer must canvass his resources and the prospects for returns in the future with a view of the best utilization of those resources. In this allocation prospective prices, naturally, occupy an important place in calculations under ordinary circumstances. With price uncertain there will be a wide variety of estimates as to the probable price and a variety of reactions to the same estimate on the part of different producers depending upon their faith in their own judgment and the inclination to take risks in the hopes of larger returns. With prices established, many of these uncertainties vanish and the calculations of what will be profitable differ materially. We will probably be wrong if we base our assumption of responses on historical relationships and do not modify them by some sort of an adjustment for this difference. This presupposes that prices are announced in advance, and such an announcement would be essential if prices are to exercise a directive force over agricultural production.

There are additional difficulties in establishing the prices of agricul-

tural products beyond determining the level for them in a particular period. These arise because of the seasonal character of production and the variation in the volume of production in various years. Many agricultural products have wide seasonal movements in their prices. These are induced by the heavy production at certain periods of the year and the necessity of storage to level out the rate of movement into consumption and for the processing of the product. It is evident that the price fixed will need to include a seasonal movement. This seasonal must be reasonably correct, otherwise great peculiarities in the rate of marketing may take place. If a single price is established for the year, then there will be no incentive to hold and everything will tend to go to the market at once unless some sort of a quota system is established for sales. Some system of monthly quotas for sellers could of course be devised, but such an administrative task would be an unwarranted undertaking in view of more important requirements for administration elsewhere. If, however, a very large seasonal advance is provided in the established price, this would lead to holding for a later period and a scarcity in the market during the early period. The choice of the seasonal must lie in the experience of the past, and some average seasonal derived from recent years will probably prove satisfactory. It is also probable that there would be an advantage in lessening the rise in the seasonal price somewhat below the average with a view of erring on the side of increasing the rate of early marketing rather than retarding it. This would insure the marketing agencies receiving the products at the full rate at which they could be adequately handled, and if only a slight price decline below the established ceiling were induced in the early period, the expected higher later price would tend to balance marketings to the more usual previous rates.

Problems are also raised by the necessity of carryovers from one year to the next. For some products these carryovers are fairly large. The carryover tends to increase when the following crop promises to be small and prices high, and the carryover declines in years when the following crop promises to be large and prices low. The price set or expected for the subsequent year will in consequence have a great deal to do with the extent of carryover during the given year. If the price set for the new crop is materially higher one would of course immediately decide to hold to the later period. Where units of new and old production are distinguishable it may be possible to establish differential prices for them to prevail in the later period. It is possible that usual marketing rates and timing may be considerably upset by failure to gauge this factor. The new situation with ceilings will differ greatly from the old where prices were free to fluctuate. Without ceilings the

outlook with respect to price in the future is uncertain, and holdovers of product may offer considerable advantages even though it appears on current information that prices will be lower in the new year. It is always possible that something will happen to greatly decrease supplies or increase demand at the last moment. Thus a small probable loss is balanced against a great possible gain if the unexpected does happen. Removal of this possibility must necessarily exercise a decided influence upon the propensity of operators to store. Instead of uncertainty and a reward for good judgment, the storage decision now becomes largely established by the price fixing. Storage either would or would not be profitable. In such a situation it may be necessary and indeed probably will be necessary for the Government to enter into the holding of products on a considerable scale and stand ready to buy or dispose of products at the established price.

This review of the problems of agricultural price fixing makes it evident that no predetermined single method will be adequate to do the job required. We cannot afford to freeze agricultural prices at any given level or in any relationship now prevailing. Furthermore, greater flexibility is required than would be obtained by tying prices to parity at any predetermined level.

Agriculture is a vast industry with many ramifications, and the price fixing must be simple and direct unless it is to become completely unwieldy and too complex for administration. The general policy must be one of utilizing the usual regulative forces of the economy to the fullest extent possible. This means that the price fixing authority must be conversant with industry and its markets and be prepared to assist the market to carry out its functions rather than to slap on a fixed set of controls. Fortunately in the accumulated knowledge and experience of the workers in agriculture, particularly in the Department of Agriculture, there is a vast reservoir to be drawn upon, and we may confidently look forward to any legislatively unhampered experiment in price fixing in agricultural products to be carried through with a considerable prospect of success.

IMPACT OF DEFENSE UPON INDUSTRIAL CAPACITY AND INVESTMENT*

By M. Joseph Meehan
Bureau of Foreign and Domestic Commerce

We are now engaged in a war in which our present productive capacity, and the additional capacity that can be brought into operation quickly are vital considerations. Therefore, an analysis of the impact of the defense program upon the flow of investment, income, and capacity provides a topic of peculiar interest at this time. An appraisal of the direction and rate of recent economic movements is essential to an understanding of the changes in the structure of industry that are taking place; what we have experienced is a forerunner of what is to come at a vastly accelerated pace as we move into the current all-out effort. I shall endeavor to sketch, in broad outline, the features of what has happened to date, and to give some indication of the implications of these developments.

First, let us take a look at what has happened since the early part of 1940. You will recall that we were then experiencing a mild reaction to the speculative surge which followed the outbreak of war in Europe. The national income in May was at an annual rate of $74 billion, with a large volume of unemployed resources available to speed up the economic machine. Once this driving force was supplied by the decision to expand armaments, it was at once obvious that new facilities had to be created for the specialized production of such goods. Not so obvious were the tremendous demands that would be made in the not distant future upon all our basic resources. Failure to secure general recognition of the latter fact led to much fruitless controversy. What was vital from the standpoint of our present position was the delay encountered in expanding those basic facilities which constitute limiting factors on our productive effort at a time when we are engaged in a world-wide struggle with powerful and resourceful enemies.

The first phase of the defense program extended over a period of a year. The immediate impact on the economy was an acceleration of the tempo of business activity as the added flow of government spending reinforced the demands from belligerents for essential commodities. Expansion at the rapid pace actually witnessed was made possible by the extensive volume of existing facilities that could be drawn into productive operation immediately. A large part of the initial outlays under the rearmament program went for the construction of new hous-

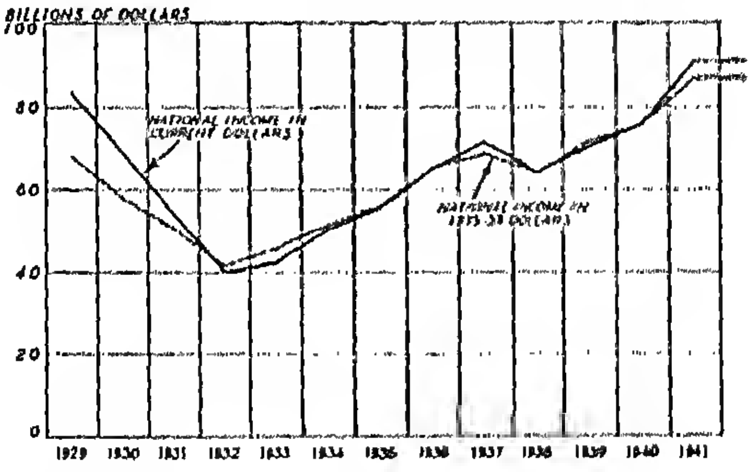
* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 27, 1941.

ing for the armed forces, for food and other quartermaster's supplies—expenditures that could be stepped up rapidly. This, coupled with the increased pressure for deliveries abroad after the fall of France, and the general change in psychology produced by the events of mid-1940, resulted in an immediate forward movement which was quickly strengthened by the initiation of the broad program of creating new capital facilities.

The result has been a rise in the national income at an unprecedented rate. Notwithstanding that the first half of 1940 was a period of some

CHART I

NATIONAL INCOME IN CURRENT AND 1935-39 DOLLARS



hesitation, for the year we had an increase in income of nearly $6 billion, or 7½ per cent. It is evident from Chart I that practically all of this rise represented an increase in real terms.

During 1941 we have experienced an advance of $16 billion, raising the aggregate for this year to an estimated total of $92 billion. It may be noted that most of this striking increase was in real terms, though the rapid expansion in physical output was accompanied by some rise in prices. In terms of 1935–39 dollars, the rise in income was from $76 billion to $88 billion.[1]

Since the basic expansion was in the industrial segment of the economy, a particularly sharp advance occurred in the output of manu-

[1] These estimates for 1941, and those given subsequently for gross capital formation, were made before full year statistics were available. Subsequent data indicate that the 1941 totals were somewhat higher.

factures which has reached record volume. The significant thing about this expansion, aside from the over-all rate, is the upward surge in the durable goods industries. When we are creating a "new" industry, or rather a complex of new industries, heavy reliance must be placed upon existing capital facilities. Before long finished armaments will represent our number one industry, and behind the finished products will be the vast organization of men and supply to turn these goods out in a hurry.

The flow of funds stemming from the arms program fanned out through the entire economy. To the demands for producers' goods on the part of the growing armament industry were added the orders from makers of civilian goods who found the time opportune for rehabilitating existing facilities or expanding them in certain instances. Activity was lifted in the producers' goods industries to unprecedented levels—more investment by far was made in new capital equipment in 1941 than in any other year in our history.

I have indicated that this phase of the mobilization of our resources extended through the first year. During that period rapid expansion was fed from existing plants and facilities, mainly through taking up the existing slack. By the middle of 1941, however, we were finding it increasingly difficult to expand production because of widespread bottlenecks in facilities. With this phase began a sharp and widespread price advance.

Additional expansion in output has been achieved during recent months, and we may expect a further rise from now on as new facilities come into production. But the significant thing to note is that we are up against limiting factors—mainly a shortage of basic materials—which were bound to restrict expansion as the tremendous magnitude of the military program made itself fully felt. We have had, therefore, in recent months curtailment in particular lines, with the area of restriction being gradually widened as military demands increasingly impinged upon the civilian sector of the economy.

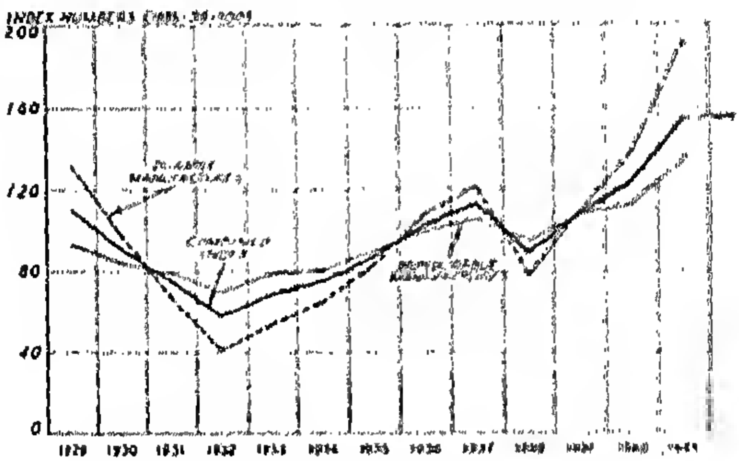
Nevertheless, while we were engaged in what was essentially the "tooling-up" process, we did add very greatly to the supply of goods available to the consumer. We enter a period of restriction of some types of goods, therefore, with larger and better-conditioned stocks than at any time in our history. The process of belt tightening which is ahead finds us well able to stand such a trend.

We may expect that, as output rises further, there will be a very marked divergence from the traditional pattern of output—a phase which is just beginning because up to now consumers have been able to secure wanted goods and thus guide, in more or less usual fashion, the production of goods. It might be noted that, even in the automobile in-

dustry, where the recent curtailment has been pronounced, we produced in the 1941 calendar year 3½ million passenger cars.

While satisfactory over-all measures of the relation of military expenditures to the total product of the economy are not available, we may accept as a relevant measure a comparison of these outlays and the gross national product. On a rough basis it is estimated that defense, or armament, expenditure as a percentage of gross product has tripled compared with the third quarter of 1940 and now represents nearly one-fifth of the total. Notwithstanding the growth of military

CHART 13
FEDERAL RESERVE INDEX OF INDUSTRIAL PRODUCTION



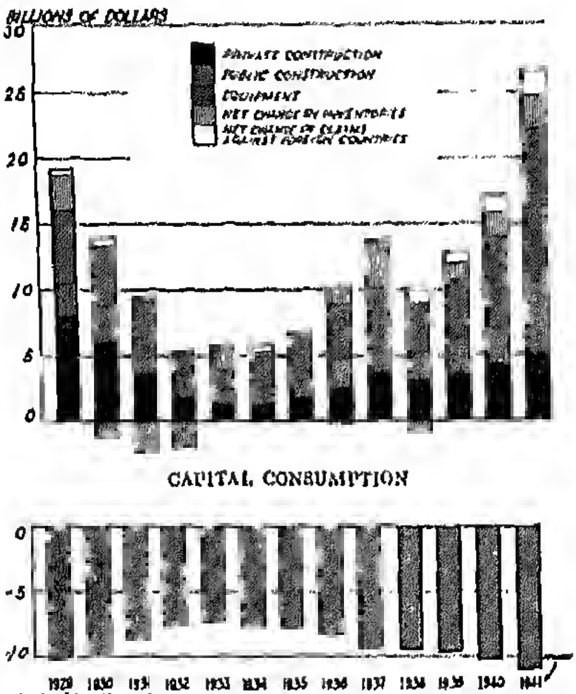
expenditures, a much larger total has still been available for nonmilitary purposes in recent months than in mid-1940 by reason of the expansion in total output previously indicated.

The proportion of military output must perforce grow. Prior to the declaration of war, we estimated that 25 per cent of the gross product in 1942 would be taken by the military program, but these estimates are at present in the course of upward revision. A huge expansion in the war program is now under way.[3] The objective will be to push expenditures as rapidly as possible towards a figure which is likely to move up towards 50 per cent of a much larger total national product.

[3] Expansion of the war program has since taken definite shape in the President's budget message January 6, 1942. The total war program covered by the message runs to nearly [illegible] billion, with war expenditure expected to reach a rate of nearly $5 billion per month before the end of the fiscal year

For purposes of analysis of investment or capital formation, the volume of capital formation has been distributed over five classifications: private construction, public construction, equipment, net changes in inventories, and the net change in claims against foreign countries (Chart III). The first three of these are always positive totals,

CHART III
GROSS CAPITAL FORMATION IN THE UNITED STATES



but the other two have been positive and negative over the period shown here. Inventories represent, of course, a particularly volatile cyclical factor.

For the national income measurement the net volume of capital formation is used—allowance being made for capital consumption. In some years over the period shown in the figure we did not make good our capital depreciation. For the present purpose, however, the series on public and private gross capital formation is the significant one. It should be pointed out that included in the figures of capital formation are the durable products acquired by the Government under the armament program.

What stands out is the magnitude of the 1941 total, and the shifting proportions of this total represented by the various categories. We experienced a rise from $10.7 billion in 1940 to $26.5 billion in 1941. These figures exclude all consumers durable goods, except housing.

Public and private construction together comprised a total of $10.5 billion. This is the same aggregate as in 1929, but the distribution between public and private construction has been radically changed, even more so than the figure indicates since some of the private construction is financed indirectly through public funds. Government expenditures account directly for $12 billion, or not far from half, of the total of $26.5 billion.

While the building of new plants is impressive, what is even more striking is the volume of equipment expenditures. Nothing remotely approaching the estimated total of $11.7 billion has previously been witnessed. Not all of this expenditure is for producers' equipment, since $3.2 billion of durable armaments is included, but the estimated $8.5 billion expenditure for producers' equipment far outstrips like outlays in any other year—it is more than 50 per cent higher than in 1929. Even the very large increase shown for inventories is dwarfed by the equipment totals.

The rise in inventories is a reflection of many factors—one of them being the necessity to stock new plants with materials and in process inventories—but it may be noted that part of it represents a store of goods which can be drawn effectively into the productive process as the armament program is accelerated and conversion of facilities takes on a more rapid tempo. This conversion will be a feature of the next year since it is obvious that further expansion of new facilities on the scale required for the war program is not possible, in view of the limitations imposed at present by materials and which will be reinforced at a later date by limitations on our man-power resources.

Expenditures on military construction and finished armaments will rise rapidly in the next year, but competing types of civilian activity will be very markedly curtailed. Likewise, we may expect the rise in income available for civilian spending to be curbed by the fiscal policy appropriate to the war effort. Investment will be conditioned almost entirely by military considerations.

This was not so true during much of 1941. While we instituted and extended a priorities system after it became apparent that many basic materials were extremely short, it is a fact that general expansion in capital expenditures proceeded along with the expansion resulting from the military program. This stemmed naturally from the high level of incomes which put increasingly large demands upon all productive facilities.

As the year progressed, however, widening priorities controls tended to curtail the flow of machinery orders and the initiation of new plant construction for civilian goods. This is reflected, for example, in new business reported by the machinery industry. From the middle of 1940 through the middle of 1941 there was a wide gap between shipments and new orders, with unfilled orders rising to record proportions. This occurred notwithstanding that the output of producers' machinery and equipment moved up from a monthly rate of $425 million in June 1940 to $700 million a year later. By October, orders had fallen to a level approximating the shipment rate which was then in excess of $800 million a month.

The slackening pace of expansion in non-defense industries in the latter months of the year held the increase in capital expenditures in these areas to a moderate rise over the comparatively high totals of 1940. Nevertheless, it is significant that at a time when the economy was under extreme pressure to expand armament facilities we added to our equipment over a broad section. This was true, for example, in such lines as food and kindred products, textiles and paper.

For the entire year 1941, we estimate expenditures for plant facilities and equipment in the manufacturing industries at more than $4 billion, or about 45 per cent of the total for all industries. In 1940, this proportion was much smaller, the $2.3 billion expenditure for manufacturing facilities representing a more usual proportion—about one-third.

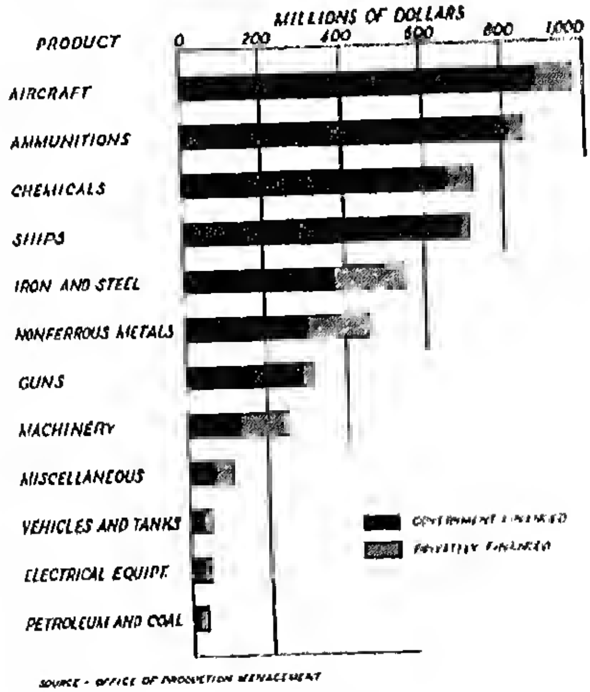
This aggregate investment in manufacturing facilities is three-fourths larger than the $2.5 billion expended in 1918 during the first World War. Today's warfare requires many times the equipment needed in 1917–18, and the facilities behind the lines correspondingly must be greater. The 1941 total is more than 80 per cent above the aggregate for the preceding year. This makes clear the magnitude of the increase in manufacturing facilities, as well as the rate at which expansion has been stepped up.

We do not have a good yardstick of total investment in manufacturing industries to use as a measure of the comparative size of the addition to our total facilities. Some indication of its size is afforded, however, by comparison with the net book value of all manufacturing plant and equipment (exclusive of land), estimated at $22.5 billion at the end of 1940. This figure, which is based upon Bureau of Internal Revenue statements of the asset value of corporations reporting for tax purposes, is no doubt an understatement of the replacement value of manufacturing facilities existing at that time. But it is easily recognized that an addition of more than $4 billion in one year, with a higher aggregate certain for 1942, represents an enormous rate of creation of new facili-

ties. We look back upon the decade of the '20's as a period of marked expansion in manufacturing capacity. During that ten-year period, capital expenditures on all manufacturing facilities amounted to $20 billion. In slightly more than two years, we shall have produced half that amount. In all the '30's expenditures aggregated $14 billion.

The distribution of the expenditure for new productive facilities by industries reveals the dominance of military demand. When we break

CHART IV
VALUE OF DEFENSE CONTRACTS FOR INDUSTRIAL FACILITIES FINANCED BY GOVERNMENT AND PRIVATE FUNDS, THROUGH SEPTEMBER 30, 1941



down the $4.2 billion total for manufacturing in 1941, we find that nearly one-half of the expenditure has been in the predominantly military industries; more than one-fourth has been in the basic industries producing for both military and civilian use; and the remaining fourth has gone into industries with predominantly civilian products.

Aircraft led the expansion of the military industries and about one-eighth of the total outlay on manufacturing plant and equipment went into such plants. Chemicals (including powder) and shipbuilding and

repair followed, each with nearly 10 per cent. Ammunition was not far below shipbuilding. The half billion dollars expended in the iron and steel industry equalled the outlay in the aircraft industry.

Expenditures in the machinery industries exceeded a quarter of a billion dollars—half of that expended in the leading industries, and not far above the totals for nonferrous metals and the fuel industry. One-third of the total expenditures in the predominantly civilian industries was made in the food and kindred products industries, one-sixth in the textile industries, and about one-tenth each in the automotive and stone, clay, and glass industries.

The distribution of defense facilities contracts indicates the general picture of the character and extent of expansion in capacity and investment that is taking place. Naturally, these data represent to a considerable extent uncompleted work.

Aircraft definitely tops this contract list with the total of facilities contracts reaching nearly $1 billion by the end of September, a figure now exceeded by the contracts placed in the final quarter of 1941. Again we need a yardstick to comprehend the size of this figure, and we have a reasonably good one in the asset returns of the Bureau of Internal Revenue. One billion dollars was the net value of the corporate assets reported by the automobile industry as of the end of 1938.

Ammunition is second on this list—well over $800 million—though it was in fifth place on the 1941 expenditure list. Chemicals and ships are ahead of iron and steel which shared the top position with aircraft in 1941. The contracts for nonferrous metals reached $400 million. Though the machinery industry is well down on the list, the total represents a very large expansion for this industry. Machine tools probably constitute the most important single bottleneck that exists, and we cannot build all the new equipment needed. We shall have to make much wider use for military purposes of existing equipment of this type. Though this is the case for all the metal working industries, the key importance of machine tools makes it especially significant in this case.

All of the aggregates here shown for military and related facilities will be expanded under the war program, but it is unlikely that the general nature of the expansion will be altered. The military needs will be such, however, that the expansion in non-military facilities on a scale witnessed in 1941 will not occur in 1942.

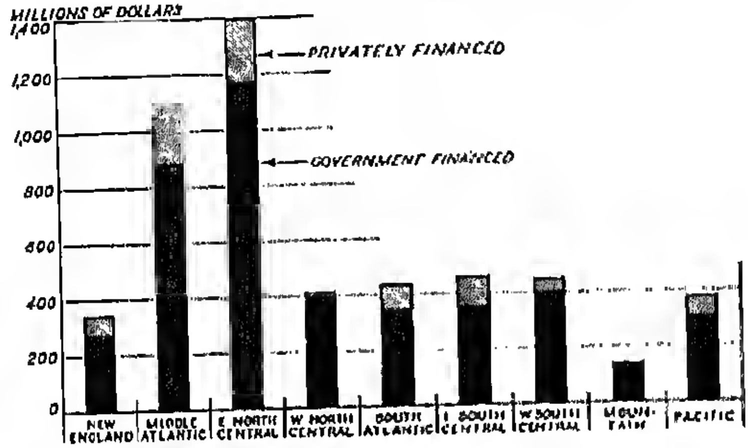
From the middle of 1940 through September 1941, commitments were made for over 3,000 plant expansions by the military authorities. When complete these enlargements are expected to cost over $5 billion. More than $4 billion, or over 80 per cent of this cost, will be met from

public funds, while less than $1 billion will be financed privately. In general, the extent of public financing is relatively greater in those industries where the output is more directly related to armament.

For example, over 96 per cent of the new shipbuilding facilities approved through September are being constructed with public funds supplied through the Navy Department and the Maritime Commission.

CHART V

VALUE OF DEFENSE CONTRACTS FOR INDUSTRIAL FACILITIES, THROUGH SEPTEMBER 30, 1941, BY GEOGRAPHIC REGIONS



SOURCE — OFFICE OF PRODUCTION MANAGEMENT

More than 90 per cent of the plant for the manufacture of ammunition, explosives, guns, and nearly 90 per cent of the new aircraft facilities are also government financed.

In contrast to this dominance of public funds in financing direct armament plant expansion, only half the new machinery plants and 40 per cent of the petroleum and coal products plants have been directly financed by the Government.

The rapid expansion of military needs has changed not only the structural, but also the geographical pattern of American industry. A rough picture of the geographical shifts occurring in defense industries may be obtained through comparing, by industries and by regions, the allocations for defense plant, and value added by manufacture in 1939. These suggest that as the armament program has developed there has been a moderate tendency for the geographical concentration of industry to decrease.

Those sections which in the past have been most heavily industrial-

ized—New England, the East, and the Great Lakes regions—while still receiving more than three-fifths of the funds allocated for new facilities under the government program, have a smaller percentage of the aggregate new plant than they possessed of the old. On the other hand, states embraced by the grain belt (the West North Central States), the Rocky Mountain region, and the South are improving their relative industrial position even though the expenditure in these areas is much smaller. Note again from Chart V the high proportion of the expansion financed by the Government.

What stands out in the foregoing review is the record addition to our productive capacity in the period preceding the outbreak of war. Further, we may note that we also added greatly to both the quality and quantity of stocks of durable goods in the hands of consumers. This latter fact is important now that we face a period in which war needs will establish the sole criterion for expansion of existing facilities. There are two reasons why we cannot rely entirely upon the creation of new productive facilities to obtain the required output of armaments. The first of these is the excessive drain that would be placed upon raw materials; the second is that we do not have time to complete such a process.

The small amount of finished material that has so far been produced testifies to the fact that our raw material shortages have been the result of the voracious appetite of new construction for these products, and not the result of a rapid rate of embodiment of raw materials in finished instruments of war. Despite record rates of construction, we confront now the major task of converting and adapting, through a process of allocation, existing facilities to war production.

The situation up to recently has been the natural result of a desire to utilize idle resources. Such idle resources may have appeared adequate for the needed *primary* expansion for war material such as airplanes, tanks, ordnance, etc. They have not been, however, adequate for the induced *secondary* expansion, for example, in transportation equipment, in homes, and in other civilian facilities, a large part of which is a function not solely of the primary expansion (as here defined) but of the general rise in incomes that has resulted from that expansion.

At present, in view of the imminence of near complete utilization of resources and the added drafts that will be made upon man power, the problem of increased investment to expand capacity for the war effort is crucial. From this point forward expansion of industrial facilities must be so spread over industry as to maximize the amount of existing resources that can be converted to the war effort. It is clear that war output must be speeded. It is equally clear that this cannot be done by

new plant expansion if the past slower rate of production has strained resources of labor, equipment, and raw materials available for plant expansion and for the building of the whole set of complementary resources that are required by new plants constructed. Therefore, a maximum amount of existing resources must be converted.

The problem is to convert the maximum possible amount of plant and equipment ranging along a scale from the completely generalized to the completely specialized or "one-purpose" units of equipment or plant. The portion of such a scale representing inadequate equipment resources, taken industry by industry, after conversion, is the amount that must be made good by the provision of new equipment. The test of its allocation, therefore, involves the determination of the use that maximizes output of existing equipment by spreading as thinly as possible the new capital resources available. It is this kind of allocation that must be contemplated for every industry considered eligible to undertake war production. Insofar as there may be any expansion of other industries this should be done in a manner to minimize the problems of meeting prospective consumer demand and of effecting the transition back to peacetime production.

THE PROBLEMS OF DETERMINING FAIR RENTS*

By Karl Borders
Office of Price Administration

In war time the formulation of fair principles of rent control for housing accommodations involves consideration of the interests of three principal parties to the rent contract: the landlord, the tenant, and the Government. Driven by the exigencies of total war, the Government cannot now permit the free operation of landlord-tenant bargaining, when the result of this bargaining contributes to inflation and impedes the smooth functioning of labor recruitment in defense centers.

Through the services of the Bureau of Labor Statistics and the Division of Research of the Work Projects Administration, rental trends are now being surveyed and charted in more than 230 defense localities. In most of these areas vacancy of rental housing accommodations has become negligible, as a result of a number of concomitant factors. (1) In-migrant workers have come to take newly-created jobs in defense industry. (2) Expanded employment has raised the economic status of many families who had previously lived in crowded quarters. Such families can now afford to "undouble" and are seeking separate accommodations. (3) Families of officers and enlisted men have moved to areas near military posts. (4) The marriage rate has risen and further increased the demand for separate dwelling units. (5) The housing stringency has caused a transfer of rental housing to the sales market, further limiting the supply of units available. (6) Because of increased costs, construction of rental housing for workers has become an even less attractive form of investment than it was before the war. (7) Priorities on housing have begun to curtail new construction in an absolute sense.

So great is the pressure of this demand for rental housing that in scores of localities, despite the exertions of private building and the allocation of defense housing, there is no prospect of anything like a free, competitive market for the duration of the emergency. The upward spurt of new construction during '39, '40 and '41 cannot be sustained. Yet the total number of dwellings built during this active period was far from adequate to compensate for the paralysis of the construction industry during the depression of the '30's. In the future we may expect the continuance of migration from non-defense to defense centers, despite efforts at the dispersion of war contracts.

For the country as a whole, the rent index of the Bureau of Labor

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 28, 1941.

Statistics for wage earners and lower-salaried workers, based on reports from 34 large cities, has shown only a moderate upward tendency: from September 1939 to October 1941 it rose only three points and even this upward movement is a reflection of increases in metropolitan centers such as Jacksonville, Buffalo, Detroit, Baltimore, Mobile, Indianapolis, and Cleveland, which are defense areas. Against this national pattern, the individual defense center presents a sharply contrasting picture.

From surveys of defense centers thus far completed, it appears that since October 1939 at least four cities with a population over 50,000 have shown an index rise of ten points or more, and 21 cities an index rise of from five to ten points in the total rent bill. From data available, there is reason to believe that, unless statutory controls are imposed, the general upward movement in these larger defense centers will continue throughout the emergency, with obvious consequences: (1) A lowered standard of living for those with comparatively fixed incomes; (2) A tendency to increase the labor turnover and encourage endless migration; (3) An impetus to demands for compensatory wage increases.

Defense localities under 50,000 had an insignificant number of vacant dwelling units at the beginning of in-migration. The large cities had some leeway. In the smaller communities defense activity brought an almost immediate and complete upheaval of the previous rental structure. This resulted in increases in the total rent bill which amounted in at least one instance to 100 per cent for the period March 1940 to September 1941. For all surveys thus far reported, the index has risen 5 to 15 points in 30 localities; 15 to 25 points in 14 localities; 25 to 35 points in 7 localities; 35 points and over in 4 localities in approximately the same period. It does not seem probable that in small defense communities an upward movement of this magnitude will continue indefinitely. The saturation point has probably been reached in some cases. After all there is a limit to what workers will pay for a broken-down shack, since they still have the ultimate choice of moving away from the area. In these communities the Government is confronted not only with the problem of stabilizing rents but of actually reducing them.

Since the need for rent control does not have to be established, the problem is primarily one of method. Whatever the factors considered, one basic principle should be determinant: The formula must have simplicity. An administrative order for rents cannot bear any resemblance to a price schedule for a highly organized industry which employs experts to deal with complicated forms. In the case of rent control, literally millions of landlords and tenants must understand their rights and duties.

The landlord's formula for determining fair rents is traditional. He

believes that he is entitled to a fair return on his investment, or to a fair return on the market value of his property. Such criteria were used in the two most significant American rent control experiments of the last post-war-period—in New York State and in Washington, D. C. The inadequacy of the results was obvious. Each piece of rental property became a potential public utility case. The administration bogged down. In an attempt to be fair to every individual, the entire program became futile.

It has also been suggested that a relationship be established between an assessor's appraisal of a rental property and a fair return. But those familiar with the pattern of municipal assessment plans and practices must realize that such a system would be built on sandy—if not swampy—ground.

In addition to fair return on investment, which implies an excursion into financial history, and fair return on a market value, which in itself is subject to inflationary influences and directly related to the rents to be controlled, there has been another proposal; the plan of those whom we may call the classifiers. They would like to establish classes of units in terms of size and character of structure and other variables, and set fair rentals for each class or category. This plan would be a delight to a real property inventory expert, but a nightmare to an administrator. So numerous are the variables in terms of age, condition and location, that any attempt to make a pattern out of the housing of an American industrial city would be doomed to failure. And even if it could be done, it would serve no purpose. A schedule of classes and rates in which the tenant does not quite know where he belongs would be useless.

Hence, by a process of elimination if not by preference, we arrive at the rent ceiling date method. This is the principle now being used by the voluntary Fair Rent Committees established under the direction of the Office of Price Administration; this method will go into effect January 1, 1942, in the District of Columbia under statutory powers; and it is basic to the rent sections of the proposed Emergency Price Control Bill. Under this method, the authority chooses a date at which rents are generally pegged for the duration of the emergency or as long as a free rental market remains inoperative, with allowance for the adjustment of hardship cases, to both landlord and tenant.

But the practice of fixing of the ceiling date is not quite so simple as the statement of the principle. Under ideal circumstances, at the very beginning of the emergency the ceiling date should have been imposed upon any locality where it was fairly obvious that the impact of the defense program was about to, or had already begun to, result in an inflationary rent movement. In most defense areas that would have

meant pegging rents in the spring or fall of 1940. The Canadians began pegging rents in selected defense localities as of January 2, 1940. We, unfortunately, have thus far had no statutory authority. Fair Rent Committees exert a psychological effect in restraining an extreme inflationary trend; they do not stop the trend, and they certainly cannot be expected to reverse a movement which is already under way. The results of this absence of statutory rent control authority are reported regularly in rent surveys.

At the present moment there is every likelihood that the Emergency Price Control Bill will include sections for the regulation of rentals in defense areas. Though the specific provisions are still before Congress, it is possible to discuss some of the problems which would be involved in setting a rent ceiling date. In fixing the ceiling date in a given defense-rental area, it would be wise to take into consideration the rents prevailing in that area on some pre-emergency date. April 1, 1940, would serve as a convenient benchmark; first, because it is the date of the Housing Census and data are available on vacancy, rent levels and type of structure for all units in the country; and second, at that time hardly any areas were yet significantly affected by the defense program, the only possible exception being a few shipbuilding centers. Having adopted a base date, it would be necessary to ascertain, so far as possible, what had happened to the operating costs of various types of housing accommodations in defense areas subsequent to that date.

From our preliminary soundings, it does not appear that in the upkeep of the dominant types of rent units in defense areas—the small one and two-family houses—there have been extraordinary increases in operating costs. There may be some—and we are tackling this problem through actual surveys of expenses from management books and through spot surveys of informed opinion where management books are not available—but there are not enough increases in general to justify a substantial advance in rentals. The facts are that in an area where the vacancy ratio is nil the landlord does not have to spend as much for upkeep as he would in a competitive market.

Thus, increased operating costs, the standard justification for most price increases, do not appear to affect rentals significantly during this period. Apartment houses which have had their fuel, repair material, and labor costs increased may show some rise, but (a) apartment houses are not a significant proportion of rental housing accommodations in most of the defense areas where problems are acute; (b) substantially increased occupancy will make up for increased costs in a unit cost analysis of upkeep expenses.

It will be necessary to continue gathering any available data on the

operating costs of rental housing from insurance companies, Government agencies, private foundations, as well as from a sampling of management books in selected defense areas, but, at best, these data will be straws in the wind. In practice it will probably be difficult to develop precise correlations between the rental movement and the upkeep costs of rental housing in a given area.

The setting of the rent ceiling date will, therefore, require administrative discretion, an analysis of each individual area, and a consideration of all the demand factors which have affected the movement of rentals in the area. It would probably be wise to set the rent ceiling date sufficiently far back to rescind as many of the exorbitant increases as may be feasible under sound administration. On the other hand, it will hardly be worth while to go back as much as two years if we peg rentals in a large metropolitan area where there has been a slow but steady rise. Among other reasons, because of tenancy changes and conversions, there would be an extraordinary number of disputes over the actual rental prevailing on the ceiling date. The administration would thus start off with controversy over the initial fact of what the fixed rental is. In small communities where the defense program has resulted in sudden profiteering increases—as indicated in an index rise of 40 or 60 points—it may be necessary to go back as far as April 1940 in order not to freeze rents at an exorbitant level.

Once a ceiling date is set for a locality, an administrator would be confronted with the problem of assuring tenants occupancy of their accommodations. Restriction of evictions is basic to any system of rent control. In the conversations which I had sometime ago with representatives of the Canadian Rentals Administration, they were very emphatic on this point. True, under a rent pegging system the landlord may not receive a rent above that in effect on the ceiling date, but if the landlord can at will order tenants to move, threat of eviction may prevent a tenant from registering complaint against a landlord violating the Act. Naturally, under any system the landlord should be able to regain possession of his property if he needs the dwelling for his own use, or if he plans a significant program of remodeling; but the tenant must be secure in his occupancy so long as he pays the legal rent and complies with other normal requirements, if he is to feel free to take advantage of his rights under the Act.

Furthermore, in any administration of rent control in a defense-rental area allowances should be made for individual hardship cases. This provision of flexbility would eliminate gross inequities. If, for example, on the ceiling date a landlord had rented to a relative at an

uneconomic rent, it is fair to allow him an upward adjustment when he rents on an economic basis. Likewise a tenant who for some peculiar reason was paying an exorbitant rent on the ceiling date should be given an opportunity to appeal for a decrease.

For units of new construction and converted units which were not rented on the ceiling date, the best administrative guide is the principle of comparability, with consideration of increased construction costs where they have occurred. What was the rent for a comparable unit in the locality on the ceiling date? This immediately raises the question as to how one determines comparability in individual cases. The only answer is that comparability in individual cases will be made on the basis of expert opinion by local informed inspectors, the same type of expert opinion which is accepted in appraisals.

Adjustments should also be made for any capital improvements effected subsequent to the ceiling date. The cost of these improvements should be pro-rated over a reasonable length of time, and the landlord should be allowed a fair addition to the rental.

Along with the freezing of rents on the ceiling date, there should be a freezing of all services which were a part of the rent contract, whether it was written or oral. Should the landlord reduce the service standard in effect on the ceiling date, the tenant ought to be given the right to appeal to the local administrator for a downward adjustment of his rental, since a reduction in services is plainly equivalent to an increase in rents.

There should also be provision for periodic raising of the rent ceiling when it can be shown that there has been a general rise in the cost of operating rental housing in a given defense area. Though at the present time there is no indication of any sharp upward movement in operating expenses, there may be in the future. After the adoption of the base date for a given area an administrator should collect all available relevant economic data, and if there are significant increases in operating and maintenance costs, he should issue an order granting the whole defense locality a specific percentage increase in rents above the base. Any such over-all allowance would, of course, necessitate an estimate of the ratio between total rent receipts and operating expenses. These ratios vary sharply from one area to another. In general, a showing of a substantial rise in over-all operating costs would be necessary before as much as a 5 or 10 per cent allowance above the rent ceiling is permitted.

Under this plan of rent pegging as of a given date, both landlord and tenant are allowed to continue the basic economic relationship which prevailed on a reasonable date after the emergency was declared. A

landlord simply would not be permitted to take undue advantage of a severe housing stringency and what amounts to a monopoly market which has resulted from defense activity.

It is obvious that any system of rent control will result in successful evasion in individual instances. (1) Because of the complete absence of rental vacancy, there may be collusion between an individual landlord and an individual tenant who may agree upon a rental above the ceiling. (2) Sales contracts may be devised which in reality amount to no more than an increase in rental. (3) There may be a general tendency to reduce services and repairs. As each of these problems arises, a wise administration will have to devise new regulations and methods to cope with attempts to circumvent the law.

On the whole, I feel that such a system of controls is likely to work out better than any previous attempts, because it is simple and because it places the burden of proof for revision upon the landlord, who has easier access to the forms and practices of administration than most tenants.

One last word on the effect of rent control upon the supply of housing in defense centers. It is an old argument that rent control will impede the progress of private construction where it is most needed. This contention becomes ever more academic as priority demands on critical materials become ever greater. However, if the only way to get new private construction is through an agreement to allow an exorbitant rental, then the job of providing housing for defense workers will have to be done by public construction of defense housing. But in my opinion a strict regulation of housing construction, so that it actually supplies the absolute needs of defense workers and not the luxury requirements of other strata in the population, can go hand in hand with a firm policy of rent control.

THE EFFECT OF THE FULL MOON ON TROUT FISHING*

By C. McC. Mottley and Daniel R. Embody
Cornell University

The situation that is to be explored, the effect of the full moon on trout fishing, is one that might easily give rise to superstitious beliefs. In approaching it one is immediately confronted with the question of selecting an appropriate method of inquiry. The scientific method is suggested as the best way of attaining knowledge and testing beliefs on any subject. Scientists—and statisticians—seem to speak as though it were a single method universally understood and used. Upon observing scientists in action, and after listening to their controversies, one might conclude, however, that different workers use different methods.

The method actually employed in a given situation seems to depend on the attitude of the individual toward his objective; many are extremely vague about it, some attempt to explain phenomena, others try to describe Nature, some seek the most precise statement of laws, many pursue Truth; not all would agree that the scientist should not only produce sound beliefs but produce them in such a way that the conclusions can be tested by the community. It is evident that many workers have never seen any need for stating their attitudes or habits of research; some would even deny any such necessity.

R. A. Fisher once made the statement that the statistician's function is to make sense out of figures. If this idea is expanded then the scientist should make sense out of such situations as the one which now confronts us. According to this view the scientist inquires into situations that do not make sense on first examination and, in so doing, he changes the state of affairs so that it does make sense. If it turns out in the long run that the community agrees with him, after trying out his conclusion, then it is judged to be a sound belief, but it is always subject to revision as more information comes to light.

The transformation of the elements of an indeterminate situation seems to proceed through a series of steps and appears to follow a fairly definite pattern. In a biometric inquiry, for instance, certain operations are needed to get from one step to the next, but the principles which control these secondary operations are usually suppressed or are taken for granted in accounts of individual research. If everyone used the same method, if, in fact, the scientific method were a single method, then taking such things for granted might be beyond question. If, on

* A paper presented at the 103rd *Annual Meeting of the American Statistical Association*, New York, December 29, 1941.

the contrary, different methods are used, and if the kind of belief produced depends on the means employed, then it becomes a matter of fundamental importance to state the method by which conclusions are reached.

The purpose of this paper is to draw attention to the necessity for keeping the directing principles that are employed in research as open to public inspection as the data, the evidence and the conclusions. In doing this the hope is entertained that it may help to do away with many of the conflicts which flare up between workers and aid in the critical scrutiny of results by the scientific community which is an essential part of the process of producing sound knowledge. The procedure will be to draw attention, briefly, to the more important operations needed to resolve an indeterminate situation and then each operation will be applied to the situation.[1]

THE INDETERMINATE SITUATION

Operation: becoming aware that known and unknown elements exist, characterizing the situation as problematic and expressing the need for inquiry.

Example: In recent years the fishing alibi has taken an astronomical trend and many explanations for poor fishing have been offered to the angling public. One of these notions is that trout fishing, even though it is carried on during the daylight hours, is adversely affected at the time of the full moon. No doubt this belief has been borrowed from marine fishermen. In the sea the effect is a real one probably related to tidal influences which are definitely associated with the phases of the moon. The existence of such a relation in freshwater appears to be doubtful. In any case inquiry seems to be needed before the idea is used to condition our trout fishing habits.

DATA

Operation: obtaining data from the situation, which is characterized as doubtful, by making observations and converting the material into symbols having both representative capacity and the power to indicate the exact problem.

Example: The idea that the full moon had an adverse effect on the fishing first became general at Paul Lake in central British Columbia after the 1934 season. In the course of fishery management studies from 1932 through 1936 catch records were kept at the single boat livery. The number of boats fishing each day and the number of rainbow trout, the

[1] Our indebtedness to John Dewey, *Logic: The Theory of Inquiry*, for the general pattern is acknowledged.

only fish in the lake, caught during the open season from May 1 through October were recorded. Over 27,000 rainbow trout representing about two-thirds of the total catch from the lake were counted. These data which are available at the Laboratory of Limnology and Fisheries, Cornell University, provide certain facts upon which the issue of the effect of the full moon may be investigated.

THE PROBLEM

Operation: establishing logical control of the situation by circumscribing and locating the particular problem presented by the data. The particular problem is: such a difference as that observed might occur by chance.

Example: A preliminary survey of these data indicated that there seemed to be a reduction in the catch at the time of the full moon in 1934 which might easily have given credence to the belief that the full moon does affect the fishing; but we want to know if there is a general effect operating consistently over a period of several years. The particular problem indicated may therefore be stated thus: was there a significant difference in the catch of rainbow trout at the time of the full moon during the six-month fishing season of the five-year period at Paul Lake?

THE POSSIBLE SOLUTION

Operation: formulating the hypothesis indicated by the data and by the exact statement of the problem. In this case chance could provide a reasonable explanation of the difference.

Example: The most plausible solution to the problem is: although variations will be found in the catch from month to month and year to year, the differences at the time of the full moon will be no greater than would occur merely by chance. In other words the difference in the catch between the two periods is not significant. If the uncontrolled variation in such situations as this (chance) forms a normal distribution then the hypothesis may be developed thus: if a hypothetically infinite population of such catch records as these gives a normal distribution with parameters, *mu* and *sigma*, then the efficient statistics, mean and standard deviation, from a population such as this one are estimates derived from a similar distribution. From these statistics it may be deduced that, if a certain variance ratio (F) is given by statistics from a normally distributed population, then the value of f to be calculated from the data is such a value as would be derived from statistics from a normal distribution. It is understood that the data are to be "normalized" in case they do not approximate the normal distribution. It is also

stipulated that the evidence is to be judged by the null principle. This stipulation converts the hypothesis into a null hypothesis.

THE DESIGN

Operation: specifying the material and the logical principles to be used in maintaining control and providing a valid test.

Example: In order to test the proposed solution of the problem, representative data were selected from the records. The week of the full moon was separated from the rest of the month; the day of, the day before and the day after the calendar date of the full moon were selected as representative of this condition. Three days were selected at random from the remaining days in each month. The number of boats fishing and the number of trout caught on each of the selected days in each of the six months in each of the five years were tabulated. The standard design for an analysis of covariance was judged to be suitable for controlling the conditions and producing the evidence, including the estimates of error, necessary for the tests of significance.[2] Since prior inquiry had shown that the mean and variance of data of this type are correlated the original data were "normalized" by transforming them to square-roots. The null principle requires, first, the prior choice of limits of significance and, second, the decision to abide by two stipulations. Since little is known about such situations as this one, it was decided that a high level of significance, the one per cent level, should be adopted. The two stipulations to be used in judging significance may be stated as follows:

1. If the value of f falls beyond the one per cent level, then both clauses of the hypothesis are denied and the hypothesis is refuted. In which case chance could not reasonably supply the sole explanation of the difference. The difference will be judged to be significant and the situation will be determinate.

2. If the value of f falls within the one per cent level, then the consequent clause is affirmed, but this does not necessarily confirm the whole hypothesis. Chance could, therefore, reasonably be offered as an explanation for the observed difference. In this case the situation is still indeterminate and further inquiry is needed in order to make final judgment, but this does not prevent an evaluation of the practical significance of the difference.

THE CONTROLLED INTERACTION

Operation: the controlled reduction of the interacting symbols which represent the data. This step is introduced because the reduced data

[2] Symbols and procedure used follow George W. Snedecor, *Statistical Methods*.

may be used to produce evidence other than that required for this particular problem; e.g., tests for differences between months or years.

Examples: The standard procedures for the analysis of covariance were put into operation for controlling and reducing the material specified in the design. Variations in the number of boats were con-

TABLE I

THE REDUCED DATA FOR THE NUMBER OF BOATS FISHING (x) AND THE NUMBER OF FISH CAUGHT (Y)

Source of variation	Degrees of freedom	Sums of squares and products		
		Sx^2	Sxy	Sy^2
Total	170	130.65036	247.63361	1129.30169
Years (A)	4	3.37352	22.74189	158.50457
Months (M)	5	10.17702	32.63908	136.95300
Period (T)	1	0.23472	0.02925	0.00365
$A \times M$	20	17.40082	55.92297	242.47799
$A \times T$	4	1.04025	3.93150	17.68021
$M \times T$	5	0.65436	2.21361	16.65734
$A \times M \times T$	20	7.20700	21.51251	101.62508
Sampling error	120	84.74887	108.03977	450.05054
Combined interactions	49	26.31633	83.68006	378.24060

trolled by the methods of regression; variations in the catch associated with the main effects, namely, the different years, the different months, the two conditions of the moon, and the various interactions were controlled by means of the independent classifications in the analysis of variance. The reduced data in the form of sums of squares and products are shown in Table I.

THE EVIDENCE

Operation: producing the required evidence from the reduced data.

Example: The evidence in the form of errors of estimate and variance ratios are shown in Table II. Since none of the interactions is signifi-

TABLE II

THE DEGREES OF FREEDOM ($d.f.$), THE ERRORS OF ESTIMATE (Sy^2), THE VARIANCES (V_y) AND THE VARIANCE RATIOS (f) DERIVED FROM THE DATA IN TABLE I

Source of variation	$d.f.$	Sy^2	V_y	f
Combined interaction	48	112.78741	2.34974	
Periods+interaction	49	114.95374		
Difference for testing	1	2.16633	2.16633	0.9219
Months+interaction	53	193.88052		
Difference for testing	5	81.09311	16.21862	7.1581
Years+interaction	52	155.96763		
Difference for testing	4	43.18022	10.79504	4.6041

cant the combined interaction term was selected as the basis for comparison.

THE TEST OF SIGNIFICANCE

Operation: the inductive inference.
Example: The value of f calculated from the data is less than the appropriate value of F in the tables of the F-distribution; the difference between the catch at the time of the full moon and the catch on days selected at random is, therefore, judged to be not significant.

WARRANTED ASSERTION

Operation: final judgment.
Example: As far as the evidence goes the samples from both periods could have been drawn from the same normal distribution. It may be concluded that this evidence does not show a significant difference in the catch of rainbow trout at the time of the full moon. According to the null principle, the possible effect of the full moon on trout fishing is still open to question.

UNIFIED SITUATION

Operation: final judgment prepares the situation for further inquiry or the criticism of the community, from which it may emerge as a contribution to knowledge. If it is judged to be a sound belief, it may then be used in further inquiry, or it may be put to direct use and enjoyment.
Example: The inquiry shows that the difference between the catch at the time of the full moon and the catch at other times, is not statistically significant. Chance variation could thus provide a reasonable explanation of the differences. The moon may have an effect on the fishing but, if so, it does not seem to be great enough to be of practical importance. If one were prepared to act on this belief, and if he had a choice of several days on which to go fishing at this particular lake, he should not let the phase of the moon influence his choice; furthermore he should not use the phase of the moon as an alibi in case he "enjoyed" poor fishing.

Space does not permit a detailed exposition of all of the principles connected with the controlled transformation of an indeterminate situation into one in which the constituent distinctions and relations are unified. In order to indicate what is meant we should like to discuss only one of the operations in more detail.

The formulation of the hypothesis is of strategic importance because, in the one direction, it represents, in the form of an explanation, the complete convergence of the relevant information contained in the

data; in the other direction, it rigorously controls the specifications to be stated in the design. The hypothesis therefore determines the validity of the test but only if it is explicitly stated in the "if parameters, then statistics" form. If the null principle is not included explicitly in the statement of the hypothesis, then it has to be understood as being a part of the attitude of the worker.

The possible solution for the problem was first stated in the form: "the difference in the catch between the two periods is not significant." Judging from the way problems and hypotheses are presented, if stated at all, in current biometrical literature, such a statement is usually considered to be adequate. If this attitude is adopted, it means that the exact test and the principles which determine the validity of the test are left to be understood, and confusion may follow. One can only trust that the investigator has been aware of the implications and is competent to handle the situation.

ESTIMATING LABOR REQUIREMENTS*

By Donald H. Davenport
U. S. Bureau of Labor Statistics

WITHIN LIMITS the total labor supply is capable of expansion or contraction. At the present time the total "normal" labor force would be about 55½ million. This number increases normally with population growth, at the rate of about 700,000 per year. However, within a short time the total can probably be expanded to well over 60 million. Young people can be persuaded to enter the labor force at an earlier age than usual; older workers can be persuaded to work beyond the time when they would usually retire; housewives may accept industrial employment on a part-time or full-time basis in addition to caring for their homes. By all of these means the effective labor force can probably be increased by as much as 10 to 15 per cent above normal levels. World War experience confirms the conclusion that under war conditions the labor force can be expanded considerably beyond its "normal" size.

This potential labor force probably exceeds the number we will be able to put to work in the immediate future. Limitations in materials and plant facilities restrict our ability to utilize our total labor supply. Even with the tremendous expansion in our armed forces which we may now anticipate there seems little likelihood that we will find our war effort limited by the size of the *total* labor force. While it seems unlikely that our *total* supply of available manpower will set limits to our war effort, there remains a tremendous problem of labor training, transfer, and adaptation. The workers who can be drawn into the labor force do not constitute a trained and experienced reserve. When men are drawn from industry into the armed forces, their replacement with inexperienced workers necessitates training and adjustment. Experience in Germany, later in Great Britain, and currently in America demonstrates that materials, productive capacity, and management bottle-necks can be progressively broken and that the ultimate limiting factor is the rapidity with which the skilled labor supply can be increased. The material and plant capacity bottle-necks allow time for the adaptation of labor resources to be worked out. Once these other bottle-necks are broken the labor shortage bottle-neck becomes acute. It is important, therefore, to look ahead and see what the prospective needs for labor are with respect to numbers, to skills, and to geographic location; and what steps can be taken to meet these needs with our labor supply.

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 29, 1941.

In this general situation estimates of labor requirements for the defense program serve two broad types of purposes:

(a) The first is to give to all groups concerned—employers, labor, and the public—a general picture of the size and importance of the problem of labor training and readjustment. After a decade of unemployment it was difficult to conceive of a situation where trained labor was to become a scarce resource. Employers, labor, and the public were slow to recognize the possibility of labor shortages and the need for pushing ahead a program of labor adaptation. The attitude of employers, labor, and the public toward prospective labor shortages tends to be a function of the immediate circumstances affecting labor supply, and changes only as circumstances change. In September 1939, when the war broke out in Europe, there were about one million fewer workers employed outside of agriculture than had been employed in 1929, notwithstanding the addition of approximately 6 million to the labor force in the intervening 10 years. Employment in all major industrial lines except shipbuilding and aircraft was below the 1937 peaks. With so many unemployed, employers could get all the workers they wanted at the factory gates. It was natural, therefore, that they were not particularly concerned with training or with the possibility of future labor shortages. By the winter of 1940 employment in the manufacture of machine tools, iron and steel products, and machine shops had exceeded 1937 employment peaks, but such important industries as steel, agricultural implements, automobiles, railroad cars, and chemicals were still experiencing employment levels below those of 1937. Until well into 1940 most industries could find all the skilled workers they needed merely by rehiring workers that had worked in those same industries in 1937. The notable exceptions were the airplane, shipbuilding, and machine tool industries. Even these industries, however, experienced no real difficulty in recruiting labor during the first half of 1940. They simply hired workers from among those who had been laid off in other industries. Under such circumstances the estimating of future labor requirements served a valuable function in getting employers to anticipate their needs.

Similarly, labor seeing considerable numbers of persons still unemployed was somewhat reluctant to accept the necessity for an expansion of apprenticeship training, and retraining programs. Information on anticipated labor requirements was of considerable assistance in securing labor's acceptance of the labor adaptation program. It also helped to impress on employers the need for developing adequate programs of apprenticeship and in-plant training. Public acceptance of the need for programs such as those indicated can best be obtained by comprehen-

sive information regarding future labor requirements as compared with the numbers of workers in various occupations still seeking employment.

(b) The second purpose of labor requirements estimates is more specific. For program planning purposes, we need to know *what kind* of workers will be needed and *where*, in order that training and recruitment policies may be planned in advance to meet the specific needs which will arise. It is useless to train workers for occupations in which the supply is already adequate to meet the need, or to train the right kind of workers in the wrong place. This suggests a further use of the estimates in connection with the planning of the location of new defense plants, and the planning of training programs for defense workers. For all of these phases of the defense program, a detailed knowledge of future labor requirements is essential to effective planning.

OVER-ALL ESTIMATES

In the early stages of the defense program, the only data available as a basis for estimates of labor requirements were the total amounts authorized by Congress for the production of military and naval equipment. On the basis of these over-all totals, it was possible to make rough estimates of the total number of man-hours or man-years of labor required to fulfill the authorized program. Thus in October 1940, the Bureau estimated that of the 10.4 billion dollars then authorized for defense, 9.3 billion dollars would be spent for manufactured goods and construction work, and that this would involve an aggregate of roughly 4.5 million man-years of labor. Until this production had been scheduled, month by month, only the roughest guess could be made as to the rate at which these man-years would be expended, and the resulting effect on the level of employment. But, taking into account the many intangible factors involved, it was estimated at that time that *defense* production would require an addition of "3 and possibly 4 million workers by the end of 1941," and that the indirect effects of this expenditure would increase non-defense employment by about 2 million.

Additional defense appropriations have materially increased the rate of expansion during the period covered by this initial estimate. The increase in defense employment during the 18-month period from June 1940 to December 1941 is now estimated at about 4.5 million. As a result of the increasing tempo of the defense program, the expansion of non-defense employment was halted in the third quarter of 1941. The increase in total employment, however, will easily fall within the range of 5 to 6 million suggested in the original estimate.

When the defense appropriations materialized in the form of defense

contracts and when detailed production schedules became available, more detailed and more accurate estimates of labor requirements became possible. But as ever-increasing appropriations were piled on top of existing authorizations, such schedules were necessarily in a continuous process of revision. And at these higher levels, the levels of production and employment became increasingly dependent not merely on the demand for goods but also on the limitations of plant capacity, materials, and trained labor. We had not then and we have not yet reached the point where these factors set absolute limits to the expansion of defense production. But the provision of additional plant, the re-allocation of raw materials, and the training and absorption of skilled labor all take time. All these factors promised to restrict, not our total production, but the rate of its increase, after the middle of 1941.

Taking these factors into consideration in June 1941 the Bureau estimated an increase in non-agricultural employment of from 2.5 to 3 million in the year ending in June 1942. This compared with an actual increase of over 3.5 million which had taken place during the preceding 12 months, when shortages of capacity, materials, and skilled labor were virtually unknown. During the first six months of the year covered by this estimate (through December 1941) the increase (affected to some degree by seasonal factors) will be about 2.2 million. As of December 1941, we see no reason to modify this estimate. It seems unlikely that total non-agricultural employment will rise by more than 0.8 million from December 1941 to June 1942.

With the rapid expansion of defense production and the curtailment of non-defense production impending, there arose a demand for estimates of "defense employment." This involved not only difficulties of estimation but also difficulties of definition. Finished products may be fairly readily allocated to "defense" or "non-defense" categories. But the total labor entering into the fabrication of any given product includes man-hours in mines, railroads, power plants, and steel mills, as well as in the plant where the final assembly takes place. And since most of this indirect labor contributes to the production of both "defense" and "non-defense" goods, its proper allocation becomes essentially a problem of national "cost accounting" in terms of employment.

Any such problem of cost accounting involves a certain degree of arbitrariness in the allocation of joint costs. Joint costs, in this case, present problems in all industries whose output is going partly to defense and partly to civilian uses, but these problems are particularly acute in such industries as transportation and power, where employment is affected much less than proportionately by increases in output. It was felt to be anomalous, for example, to allocate any substantial

proportion of the employment in the transportation and public utility group to defense at a time when employment in this field had increased only slightly, despite a considerable apparent increase in both defense and non-defense "output."

Thus certain more or less arbitrary "cost accounting" conventions were adapted in defining "defense employment." Physical output in each industry group is first broken down into its defense and non-defense components. Employment in a "pre-defense" period (we have used the second quarter of 1940) is allocated proportionately to defense and non-defense categories on the basis of the corresponding allocation of production. The employment and production data for each quarter are then compared with those for the pre-defense "base" quarter. If non-defense production is above the base level, the employment *increase* (if any) is allocated proportionately to the production increase to the defense and non-defense categories. If, however, non-defense production is below the base quarter, non-defense employment is regarded as having declined proportionately, and the remaining employment allocated to the "defense" category. On this basis, an estimate of defense employment for the fourth quarter of 1941 will be about 5 million, as compared with roughly 0.5 million in the second quarter of 1940.

SPECIFIC ESTIMATES

Such estimates as have been described, however, do little more than indicate the general magnitudes of the problems we face. While they have value in connection with broad policy problems, the detailed planning of training, housing, or production programs requires more specific estimates. Therefore, more than a year ago we undertook to prepare specific estimates of labor requirements for the principal defense industries.

A substantial part of the early defense program was made up of construction projects—cantonments for our expanding army; new facilities for the Army and the Navy; and new industrial facilities for increasing our output of plants, guns, and tanks. The Bureau's accumulation of data on PWA construction projects made possible the prompt preparation of estimates of requirements for construction labor.

Similarly, estimates of the labor required in connection with the production of aircraft and ships were prepared immediately after the first contracts were let, and have been revised with each major expansion of the program. In this case, since large parts of the program were placed under contract very promptly, it was possible to indicate quite accurately not only the total volume of labor required from month to

month, but also its geographic distribution. Studies of the occupational patterns of labor in shipyards and aircraft plants made it possible to go even further and indicate the occupational composition of the increased demand.

For ordnance items such as tanks and guns there were no basic data in the Bureau's files which would indicate the number of men or man-hours required. It was necessary therefore to obtain from the manufacturers and the arsenals estimates of the labor they were going to require to complete their contracts.

These specialized estimates cover only the most important segments of the entire program. They do not add up to the total estimates described earlier, nor should they. However, they have served an extremely useful purpose—that of supplying detailed specific data on the critical defense industries.

THE IMPACT OF WAR ON LABOR SUPPLY AND LABOR UTILIZATION*

By Louis Levine
Social Security Board

THE EXPERIENCE of the past 18 months of the defense program with respect to problems of labor supply and utilization of labor will stand us in good stead during the war period. From a decade characterized by huge surpluses of workers willing and able to work but unable to find a job, this Nation has moved into an era in which acute stringencies of labor in many occupations are already widespread. Shortly after the rearmament program got under way, shortages of skilled workers in a number of highly essential occupations required by important defense industries confronted us. The shortages in machine tools and aircraft dated back to the fall of 1939, when the British and French demands came into being. The requirements of our own defense program accentuated the labor supply problems in these industries and quickly brought to our attention the needs in such other industries as shipbuilding and ordnance. The most serious aspect of the labor supply problem is that of obtaining the manpower required for a relatively small number of skilled defense occupations. During the defense period, it was difficult to awaken employer concern regarding effective utilization of labor resources. Furthermore, as the impact of shortages of critical materials and curtailment of civilian production came to the fore, there was a tendency to be less concerned about availability of labor and to give increasing attention to so-called priority unemployment. The tremendous expansion of armament production now required once again brings into the limelight the serious problems of labor supply and efficient utilization of labor.

Modern warfare differs from earlier wars chiefly in the degree of mechanization of the armed services and, therefore, the extent to which skill is required both in civilian and military groups. The production and replacement of aircraft, ships, tanks, and other ordnance requires a higher ratio of skilled workers to total employment than has been necessary in the past. The far-flung theaters of war and the intensity of war operation place a tremendous burden on the civilian population to keep the armed services supplied. The experience of the past 18 months is but a foretaste of the problems which now arise and the recognition of the gravity of these problems recently has resulted in the taking of several important steps. The nationalization of the public employment

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 27, 1941.

offices into the presently constituted United States Employment Service represents a recognition of the national character of the labor supply problem and provides an important mechanism for the organization of the labor market and the orderly occupational and geographic transfer of workers from civilian to war-time production. The recently enacted legislation for registration of all males between the ages of 18 and 64 is directed toward the mobilization of manpower, not only for military service, but also for civilian production to support the armed service.

It became evident quite early that the problems of labor supply and efficient utilization of labor could not be resolved without current and comprehensive labor market information upon which basic policies could be formulated. In contrast to the situation faced by this country in the last war, we have today an extremely important body of labor market data, and the resources of the local offices of the United States Employment Service bring forth currently from each locality an indication of the supply and the demand for labor as well as the changing characteristics of the labor market. Each month the public employment offices report the number of available qualified workers registered in over 900 of the occupations most essential for the war program. Each 60-day period representatives of the employment offices visit employers in selected war industries employing 50 or more workers to obtain their prospective hiring and lay-off schedules by specific occupations for the ensuing 6-month period. In November, some 12,500 plants employing approximately 5,700,000 workers reported their prospective employment schedules. Each month the public employment offices report for each locality and by industry the important developments in the labor market with respect to changes in hiring practices and specifications, scope of the defense training programs, migration, housing problems, etc. The United States Employment Service, through its 1,500 full-time local offices and 3,000 part-time offices, has become the central source of labor market information for other governmental agencies concerned with labor problems. For example, approximately 300 local labor market surveys have been completed. These surveys intensively analyze the economic characteristics of a labor market area, the current and future demands for labor, the size and composition of the available labor supply, the methods for recruitment of labor not available locally, and the general community problems affecting the labor market. Virtually every important community in the country has been surveyed and some of these have been re-surveyed many times. Such information is extremely important in determining the allocation of defense housing, the awarding of armament contracts,

the location of war plants, and the orientation of training programs. As material shortages and civilian production curtailment began affecting increasingly large numbers of industries, public employment offices obtained information from individual employers affected regarding the number and types of workers laid off or to be laid off in order that such workers might be utilized by other employers for armament production in the event that plant conversion by present employers could not be effected. Where an entire community was threatened with serious unemployment because of the rearmament program, local community surveys were undertaken by the employment offices and in some instances such communities have been certified as distress areas, thus permitting employers in these areas to be given special consideration in the awarding of armament contracts.

Obviously labor market information, important as it is to the formulation of labor supply policy, is only the beginning in meeting the problems of labor supply and utilization of labor. The defense program, and now the war effort, exerts a vital influence upon the size of labor supply, its composition, and its utilization. At all times entrance to and exits from the labor market are extremely important determinants of labor supply. The increased earning power and the large volume of job opportunities have influenced many young persons to enter the labor market. Likewise, retirements from the labor market have been reduced and many workers who had already retired have returned to employment. The character of employer specifications is also an important factor in influencing the labor supply. The increasing shortage of skilled workers has already been reflected in the relaxation of employer specifications with respect to age. There are, however, many restrictive specifications which interfere with fullest utilization of labor supply. Restrictions against the hiring of non-citizen workers prevail in almost all industries, although legal limitations apply only to employment in the aircraft industry and to work on secret contrivances. Similarly, there is widespread evidence of a failure to hire Negro workers. During the third quarter of 1941 the public employment offices made 149,000 placements in some 20 selected defense industries, but only 3 per cent of such placements were of non-white workers. The greatest gains in Negro placements have taken place in the service occupations. There is, however, a tendency slowly emerging to hire Negroes in such occupations as cement finisher, bench molder, and form builder and shaper. Thus far the entrance of women into the labor market has largely taken the form of replacement of men who have left the service and trade industries to enter manufacturing. Only recently has the tendency emerged to hire women as production workers. In aircraft, women

workers are now used to assemble airplane radios and as polishers and aluminum welders. In some industries, increasing numbers of women are being used as spot welders, tracers, solderers, and milling-machine and drill-press operators. There can be no question that the most efficient utilization of our labor will not take place until restrictive specifications unrelated to the occupational qualifications involved are removed.

An important source for augmenting our labor supply lies in our training program. The unfortunate experience of the 10 years preceding the defense program during which apprenticeship and in-plant training activities were discarded is difficult to overcome. Great strides have been made, however, in the expansion of the defense vocational education program. Over 1,775,000 persons were trained during the past year and 50 per cent of such training was in essential occupations. The demand for labor today, however, indicates that employers are frequently in such need of labor that they cannot wait until the completion of training courses. The defense vocational training program does not turn out highly experienced or skilled workers. The training period is short but if workers are to be made available more quickly, training periods will probably be shortened even further. Increasing emphasis will have to be given to in-plant training and perhaps less to pre-employment training. As job processes are scrutinized more carefully and as such processes are simplified into their component parts, less-skilled workers will be required. To a considerable extent the shortages in the highly skilled occupations can be met only by upgrading of workers from less-skilled occupations and by transfer of skilled workers from civilian to armament production. We do not now have in our labor supply unemployed skilled workers in the essential armament occupations. The effective utilization of labor requires occupational and worker analysis. The work of the occupational analyst in the United States Employment Service, the availability of the occupational dictionary (covering some 18,000 occupations and prepared by that service) and the increasing awareness of management with respect to the possibilities that lie in occupational analysis, are important factors in augmenting our labor supply.

In normal times the armed forces of our country do not have an appreciable influence upon the labor supply. Even with an Army of a million and a half young men, our labor market was not seriously affected. With a prospect, however, of 4 to 7 million men in our armed services, the problem of labor supply becomes increasingly serious. It is extremely important that the needs of our armed services be met, but it is equally important that the production of supplies for our armed

services not be impaired. For this reason, the selective service process must give increasing attention to the problem of occupational deferment. The final test lies in whether a man will contribute more to the war effort in an armament plant or in the armed services.

There are, of course, many other factors that influence the labor market such as wage rates, hours of work, and adequate housing. These factors I conceive to go beyond the scope of this paper. They are nevertheless important, for they impede our war effort to the extent that they disrupt production, create labor disputes, and cause high labor turnover. There are other factors such as costs of living and rent which also have a bearing on the labor market. To discuss all of these would require more space than limitations permit.

No discussion of labor supply and utilization of labor would be complete without some attention to the problems of distribution of defense contracts by industry and geographic area. Certain sections of the country have been overburdened with armament work to such an extent that serious labor supply problems have emerged and in their wake have come problems of housing, sanitation, schools, etc. On the other hand, there are other areas in the country which have been affected only slightly by contract awards. The concentration of defense contracts in certain areas along the Pacific Coast and in various sections of New England, Middle Atlantic, and Great Lake States has stimulated a vast migration of workers from other parts of the country. The maldistribution of contracts has undoubtedly accentuated the problems of labor supply and dislocation. The strenuous efforts now being made to spread contracts more evenly throughout the country become more important than ever. Sub-contracting must be increased to the point where every plant and facility, no matter what its size, may participate in the war program.

Available data indicate that wide variations exist throughout the country in the distribution of defense contracts for other than construction and the supplies of labor available for employment. The New England States, for example, have approximately 12 per cent of the contracts and defense employment, but only 5 per cent of the active file of persons seeking work through the United States Employment Service; nevertheless, about 8 per cent of the hires defense industries expect to make by May 1942 will take place in New England. Even more strikingly, the Middle Atlantic States have approximately 32 per cent of the contracts and defense employment, and employers in these states anticipate making nearly 30 per cent of the hires in defense industries anticipated by May 1942; these states have only 21 per cent of the active file. The situation in the South and the West is far different. The

Southeast States, for example, have received only 5 per cent of the contracts, have only 3 per cent of current employment in defense industries, and anticipate making about 7 per cent of the defense hires between now and May 1; they have over 12 per cent of the active file. Likewise, the Gulf States have only 4 per cent of the defense contracts, 3 per cent of the current employment and anticipate making only 6 per cent of the new defense hires through May 1; they have about 14 per cent of the active file.

Perhaps the best approach to an analysis of the impact of the war program upon labor supply and utilization of labor is to proceed to a study of specific labor market areas, each of which illustrates a different type of situation in which adjustments in labor supply and utilization of labor became necessary.

Detroit. Detroit at the moment is our most spectacular illustration of the labor market dislocations caused by the outbreak of war. Virtually every problem associated with a shift from peace-time to war production is reflected in this area. From our most important urban center for the production of consumer durable goods, the city, because of the necessity of war, is being transformed into a great arsenal of planes, tanks, and guns.

At the beginning of 1940, the Greater Detroit area had a population of over 1,800,000, of whom about 360,000 were gainfully employed. By the end of 1941, employment had increased 17 per cent to a total of 411,000 workers, of whom about one-third were engaged in defense production. The additional workers employed have been recruited both from the locally available unemployed reserve and from the large numbers of workers migrating to Detroit in response to job opportunities and high wage levels.

The outlook for 1942 forms a rather confused picture. Industry primarily the automobile industry affected by drastic curtailment of automobile production, must convert existing facilities and increase defense production at a far more rapid rate. During the transition period, thousands of automobile workers will be thrown out of employment, with little immediate prospect of reemployment on defense work. The estimated number that will be displaced during the early part of 1942 varies between 50,000 and 100,000 workers. Reemployment depends, to a large degree, on the speed with which plants can be converted and on the completion of new facilities. There are, in addition, many complicating factors, such as shortages of certain types of skilled workers and the necessity for large-scale training and retraining of workers in order to fit them for the new jobs. It appears likely that absorption of the displaced automobile workers will proceed rapidly

during the second half of 1942 as the output of planes and tanks grows to large proportions. An actual labor shortage may develop, particularly if displaced workers migrate to other areas in search of employment.

New Orleans. New Orleans illustrates an urban, non-industrial city which almost overnight has become one of our chief producers of direct war tools, such as cargo ships and fast torpedo boats. In 1940 New Orleans, a city of half a million population, was principally a transportation terminal, the major port of the Mississippi Valley and coordinating center for railway, motor trucks, airway, inland waterway, and deep sea steamship services. Of a labor force comprising more than 200,000 workers, only 20,000 were employed in the city's manufacturing industries. Of these, over three-fourths were engaged in food-processing and in the production of textiles and clothing. Shipbuilding and repairs furnished employment for only 4,500 of the community's gainful workers.

By September 1941 the number of workers in manufacturing employments had increased by 45 per cent over the previous year, and were turning out more than 20 different products to fulfill primary defense contracts totalling about 43 million dollars. In addition to the great increase in shipbuilding and ship repair activities, defense contracts have expanded the iron and steel, machine shop, and clothing industries. Increased production has required the expansion of existing facilities as well as the building of new plants and shipways.

Recent estimates put the number of additional workers to be hired by September 1942 at 8,000, nearly all to be employed by the shipbuilding industry. The labor requirements of defense production have been and will probably continue to be met largely by the locally available supply of unemployed workers. In-migration is limited to a few key technical workers.

Philadelphia. Philadelphia is an example of an urban center making products indirectly related to defense needs which has been transformed into one of the Nation's most important defense centers. The city began declining as an industrial center in the early 1920's. By the time the 1929 depression hit the country, the demand for its industrial products was so drastically reduced that the city became one of the Nation's most seriously depressed communities. The 1940 census showed Philadelphia as one of the cities which experienced a decrease in population during the last decade.

But from June 1940 to October 1941 the city has received War and Navy Department contracts, especially for shipbuilding, metal products and textiles, amounting to approximately $1,000,000,000. The

effect of this deluge of war orders became evident everywhere. Idle factory space, which seemed a liability a year and a half ago, is now at a premium. Dismantled shipways and those in disrepair are being rebuilt and reconditioned. At present, machinery, long idle in transportation equipment factories, is making guns, tanks and ammunition. Shipyards are roaring with work on rush orders.

The local supply of labor is far too inadequate to meet demand. Thousands of workers will have to be brought into the area. However, near by are areas of surplus labor such as the coal mining section of eastern Pennsylvania, and the agricultural areas of eastern Pennsylvania, southern New Jersey and Delaware, which can be easily tapped by the city.

Radford-Dublin. What has happened in the Radford-Dublin area of Virginia illustrates the impact of defense upon a rural economy into which a major war plant is suddenly placed. These small towns are the centers of a constantly expanding powder manufacturing and bag-loading plant, respectively. Previously, considerably more than one-third of the 46,000 gainful workers in this labor market area followed agricultural pursuits and only 9,000 workers were engaged in the manufacturing processes.

Within the relatively short space of one year, the communities and their environs were transformed from a primarily agricultural and domestic economy to a refurbished industrial "arsenal for democracy." By December 1941, the number of workers in manufacturing had almost doubled to approximately 17,000, radically changing the industrial composition of the area.

The labor supply available to meet these extraordinary demands was estimated at 10,000, of which half were men formerly employed on the construction of the plants. Most of the available workers were qualified only for unskilled and semiskilled occupations.

Hiring techniques involved on-the-job training, upgrading, and the transfer of key workers from other key plants of the operating company. An outstanding source of supply was the use of women, more than 2,000 of whom were employed at the bag-loading and textile plants alone. Finally, wage rates were sufficiently above those prevailing in the area to enable the management to compete successfully for workers, despite the danger involved in handling powder. Employment of Negroes presented no problem because of the predominantly white character of the population.

The development of the Radford-Dublin area illustrates effective utilization of America's labor supply. The relatively simple productive processes of the two plants permitted the tapping of a large labor pool

without the inevitable dislocations arising from a heavy influx of workers.

A review of our labor supply problems and the utilization of labor as currently evidenced in our war industries indicates that we are not likely to be faced in the very near future with a general labor shortage —that is, a shortage of manpower and womanpower.

However, this conclusion should give rise to no complacency in the formulation of basic labor supply policies. The immediate task of adjusting men to jobs and jobs to men is a formidable one. The transition from civilian to war production has not yet been made and many serious problems involving geographical and occupational transfers lie immediately ahead. We cannot afford to tolerate the labor market malpractices now emerging, such as labor scouting and piracy, unrestricted advertising, and competitive dislocations. Neither can we afford the aimless migration of large numbers of workers in response to often ill-founded rumors of job opportunities. The responsibility of the Government with respect to organization of the labor market is greater now than ever before. Only through the formulation of labor supply policy by Government, closely geared to the schedule of armament production and to the changing needs of military operations, can our war effort be brought to a successful termination. The challenge is great, but the resources of Government, management, and labor in this country are equal to the task.

ON CERTAIN BIASES IN SAMPLES OF HUMAN POPULATIONS*

By Jerome Cornfield
U. S. Bureau of Labor Statistics

Historically, samples are an outgrowth of case studies, not of the mathematical theory of sampling. The case study derived conclusions on the nature of a parent population on the basis of the characteristics of a single individual. The sample was originally a collection of case studies. Thus in the field of family expenditure studies one of the pioneer works was LePlay's six volume work, *European Workers*, based upon an intensive study of the incomes, expenditures and savings of 36 families. It would not be unfair to say that up to the present day European studies of family expenditures have been extensions of the technique developed by LePlay rather than applications of the mathematical theory of sampling.

Now the data yielded by mere collections of cases are not without value; they may give the investigator greater insight into his subject; they may strengthen his intuitive grasp; they may possess a high order of intrinsic interest. The one characteristic they lack, however, is the ability to provide statements of a known order of accuracy about a parent population.

The best known example of a collection of cases which was completely incapable of performing the function of a sample, providing an estimate of the value of a population parameter, is, of course, the 1936 *Literary Digest* poll. In the issue immediately after the election returns had been received the *Digest* commented:

> Other statisticians took our figures and so weighted, compensated, balanced, adjusted and interpreted them that they showed Roosevelt. We did not attempt to interpret the figures, because we had no stake in the result other than to wish to preserve our well-earned reputation for scrupulous bookkeeping.

The lesson of the *Digest* poll is less, I think, the fact that telephone directories and automobile registrations will not furnish an unbiased estimate, no matter what the size of the sample, than that a sample is not a matter of scrupulous bookkeeping. As a matter of fact, it was possible to take the published *Digest* results and estimate a Roosevelt majority. The poll showed that 52 per cent of the persons returning ballots in 1936 who also reported their 1932 vote had voted for Hoover in 1932, whereas Hoover had actually received 41 per cent of the 1932

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 28, 1941.

vote. By reweighting in such a fashion as to assume that the poll was correct in showing the way in which 1932 Hoover and Roosevelt voters would vote in 1936, but incorrect in showing the relative proportion of Hoover and Roosevelt voters in 1932, an approximately even split in the vote between the two major candidates was indicated. Since this was the most favorable assumption that could have been made in view of the *Digest's* sampling procedure, a Roosevelt majority was clearly indicated. At least as important a source of the *Digest's* error as the selection of the sample therefore was its failure to go beyond matters of scrupulous bookkeeping.

A sample of families obtained by drawing names from a payroll listing earners may be something worse than a collection of cases, if the data obtained from them are not suitably handled. It may not be entirely obvious that such a procedure has a serious bias. The chance of a family being included in the sample becomes dependent upon the number of earners it has; the greater the number of earners the greater the chance of its being included. Since families with several earners generally have higher incomes, more adult members, and fewer children, this sampling procedure would result in definitely biased estimates, had the sample been regarded merely as a collection of cases. A sample drawn on this basis may be used to produce valid results, however, if it is properly weighted. This problem came up in connection with the Bureau of Labor Statistics' study of the family expenditures of wage earners and clerical workers.

If p per cent of all workers are included in the payroll sample, then it is easy to see that the probability of a family with i earners being included in the sample is $(1-q^i)$, where $q=1-p$. If there are N_i families with i earners in the parent population, the expected number of families with i earners given by payroll sampling is $(1-q^i)N_i$, whereas the number given by an unbiased procedure would be pN_i. The procedure devised by the Bureau to counteract this, in effect, involved grouping families by the number of earners and assigning to each group a weight. The weight for the ith group is $p/(1-q^i)$. For a sufficiently small p this expression reduces to $1/i$, the weighting factor that would intuitively occur to one.

An alternative method of handling the problem, of course, is to include the family in the sample, only if the sampled earner is the chief earner, or more generally has any characteristic which only one earner in the family can have. This same problem arises, of course, in the selection of a sample of families from census enumeration sheets listing each member of the family. It has also occurred in drawing a sample of law firms from a directory of lawyers.

There are probably two important obstacles to drawing samples from finite populations which can be used to make predictions. First of all, to draw such a sample we require what amounts to a complete listing of the units in the parent population. If there is no complete listing, there is no possibility of devising a sampling procedure which gives each possible sample the same chance of being drawn. And, if this condition is not satisfied, the order of accuracy of any inference about the parent population is unknown.

Many expedients have been used to derive such lists but it is doubtful whether any of them has been entirely successful. The city directory has been widely used to draw a sample of households and yet a sample count of the households listed in the 1941 Washington City Directory indicates that it contains 25 per cent fewer households than the Census enumerated in April 1940.

When a sample of individuals or households is taken it is usually possible to dispense with a complete listing of the parent population by taking a sample of areas, when a complete list of the areas within which that population is located is available. After that, one can either enumerate all individuals living within the sampled areas, the procedure usually followed when simple schedule forms are used, or alternatively prepare lists of all persons living in the sampled areas and sub-sample from these lists. The latter is the procedure employed by the Works Progress Administration in its labor market studies.

When the parent population being sampled is relatively small, however, but dispersed over a large area, the cost of areal sampling may be well-nigh prohibitive. Such is the case when a sample of industrial establishments, particularly within a single industry, is taken. The absence of a listing of the parent population in such a case may be illustrated by the sample involved in indexes of employment for industries with a large number of small establishments. Changes in employment are a result of two factors: changes in the average number of employees per plant and changes in the number of plants. To estimate the first factor it would be sufficient to draw a single sample of plants and compute the number of employees in them in successive months. To estimate the second with any precision, however, it would be necessary to draw a new sample of plants every period. To do this, however, would require a new directory of plants in each industry every month. As a result, estimates of changes in employment in many industries are essentially estimates of the first factor only. Over a period of several years, in particular, such estimates frequently develop a marked cumulative error. To correct this it is necessary to adjust the estimates to the results of the biennial Census of Manufactures after making certain

that the industrial classifications are identical. Given the present state of industrial directories, it is doubtful if anything more could be done to improve the estimates, at least on a sample basis.

A second serious source of error in field surveys is the refusal rate. Human populations differ from balls in an urn at least to the extent that the balls cannot object to being included in the sample. The most usual, but certainly not the only case, in which such refusals occur is in the survey conducted by mailed questionnaire. Frequently the questionnaires are mailed to a substantial part of the parent population and whatever fraction of the original questionnaires are returned are tabulated, and the results published as a sample.

A minimum requirement for such surveys is to compare some characteristics of the parent population with the characteristics shown on the returned questionnaires. In a recent survey of the housing conditions of federal employees, for example, the necessity for prompt release of results dictated the use of questionnaires distributed directly to a sample of federal employees at their desks. It was possible to predict at least one result in advance. Employees on leave or in the field would not return questionnaires. The amount of bias that could be introduced by this factor alone was unknown. To obtain some indication of it the sample of employees was drawn from Treasury payrolls and both the name and the monthly earnings of each employee so sampled were recorded. The income distribution shown by the returned questionnaires was compared with the distribution given by the sampling from the Treasury payrolls. The amount of bias, at least with respect to income, was comparatively small. Roughly 75 per cent of all questionnaires distributed had been returned, with slightly more than 75 per cent at the lower income levels and 61 or 62 per cent at the two highest. The possibility of substantial biases with respect to other factors still existed, however. This relatively small differential refusal rate is probably unusual for questionnaires. It is at least partially explained by the fact that the questionnaires were distributed by the administrative officers of the various agencies, and were filled out in office time. More frequently, when such an analysis discloses biases, or more precisely differential refusal rates with respect to different factors, it is necessary to weight the sample by small categories or strata to compensate for the refusals. More generally, attention is now being given to the possibility of an actual field interview of a sample of the non-replying persons. This technique, if properly handled, may give unbiased estimates while retaining much of the economy of a mailed questionnaire.

Even in field interviews, as distinguished from mail surveys, the refusal rate may seriously bias the returns. First of all, sampled persons

may not be at home at the time of the agent's visit. More frequently, however, there is a minority of the population that may refuse to give any information. The Consumer Purchases Study, for example, had been sufficiently identified in the minds of many people with the New Deal for a substantial refusal rate in the upper income groups to appear. Despite the attempt to substitute families of the same economic status for the refusing families, this bias was still marked. Thus out of 178,000 income tax returns filed in Chicago, almost 10,000 were for incomes of $7,500 or more, whereas the estimate of the number of Chicago families with incomes of this amount made on the basis of the Consumer Purchases Study sample was less than 4,000. This discrepancy is, of course, too large to be explained by sampling error. In preparing estimates of the income distribution of all families in the country on the basis of the Consumer Purchases Study, the National Resources Committee found it necessary to supplement the sample data by the information afforded by income tax returns.

In some cases the refusal rate may be definitely affected by the type of schedule used. In collecting data on family expenditures, for example, there have been two generally accepted methods. One has been to send an agent to each sampled family and ask a responsible family member to recall in detail the family expenditures during some period preceding the agent's visit. The other has been to ask each sampled family to keep a detailed statement of its expenditures as they are made. The merits of the methods have been discussed pro and con for some time, with the first method in general use in this country and the second in general use in Europe.

In its 1934–36 survey of the family expenditures of wage earners and clerical workers, the Bureau of Labor Statistics experimented with both methods. The bulk of its data was collected by means of a schedule which relied upon the housewife's memory. Partly for the purpose of checking on the accuracy of these reports a certain per cent of the families were revisited and asked to keep detailed records of their food purchases and consumption. All food in the larder at the time of the agent's first visit was weighed, the housewife was furnished with a set of scales to weigh all purchases, and the food on hand was weighed again at the end of the reporting period.

Now it is apparent that families willing to undergo such an experience may not be completely representative of all families. A comparison of the families who kept these detailed records with the families who gave schedules based upon memory but did not keep detailed records discloses several characteristic differences between the two groups. Out of the approximately 1,500 families in the 4 cities included

in this experiment, about 300 kept records of their food consumption for one week, 1,200 did not. Of those keeping records 10 per cent had home makers employed outside the home; of those not keeping records, 20 per cent had employed home makers. One of these research methods thus had the effect of causing a much higher refusal rate among families with employed home makers than the other. Significantly different refusal rates were also discovered with respect to type of family. Families consisting of husband and wife only, and families lacking at least one spouse kept records much less frequently than families with husband, wife and other members.

Perhaps an even more interesting consequence of keeping records is the actual effect of the keeping of the record upon the consumption of the family. The purpose of recording food on hand at the beginning and end of the reporting period was to obtain an accurate measure of consumption, as distinguished from purchase. On computing the value of inventory change, however, we discovered that virtually all the families had less food on hand at the end of the reporting period than at the beginning. This tendency was observed in every city studied and for virtually every food analyzed. Apparently to avoid weighing each purchase almost all the housewives had discovered the expedient of consuming as much as possible from their larders. Estimates of the food purchases of families were consequently biased downward. In fact a comparison of the purchases of the same families as shown on the schedules which relied upon memory, and records kept for a period not more than 30 days distant from the period to which the schedules referred, showed that total value of purchases was significantly lower when records were kept than when memory was relied upon, while the value of food consumed was significantly higher.[1]

Examples of bias in field surveys can be multiplied indefinitely. In general, we may say that the possibilities of bias are always present in even the most well-conducted of surveys and are frequently beyond control. Consequently they must always be explored and the possible necessity of weighting to reduce the resulting bias constantly kept in mind.

[1] A more complete analysis of this experiment is now being prepared.

DEFENSE MIGRATION AND LABOR SUPPLY*

By Howard B. Myers
Work Projects Administration

For the second time in a decade migration is front page news. Again feature writers are presenting the life histories of migrant workers and their families, camera men are busily photographing the "typical migrant," newspapers and popular magazines are describing the frenzied conditions in destination towns. Again the Nation is hearing the story of "America on the move."

There are important differences, however, between the migration of the early thirties and the present movement. This is boom, not depression migration. Whereas in the thirties migration was stimulated by sharp employment declines, now it is stimulated by equally sharp employment gains. Whereas in the thirties most migrants were poverty-stricken, unsuccessful in their search for steady work that would yield a decent income, most migrants today appear to be finding jobs at relatively high wages. Whereas in the thirties migration was largely negative—an attempt to escape from known misery—present migration is largely positive—a move with the expectation of steadier employment, better wages, or the acquisition of a higher skill. And, whereas in the thirties, migration was viewed with serious concern by the public, today it is received with general public approval.

In a broader sense, of course, the stimulus of both defense and depression migration is the same. The drive that caused the large scale westward movements of the 19th century is still operating. From the point of view of the migrant, migration is an expedient resorted to in an effort to improve his economic status. In most cases he moves in the hope of finding employment, getting a better job, or obtaining higher wages.

The difference in public reaction to present-day migration results from the fact that, after 10 years of large-scale unemployment, the Nation is suddenly experiencing a boom market for labor. Unemployment is still large, but it is assuming less formidable proportions. Faced by a sharply growing demand for workers in war industries and by large-scale withdrawals to the armed forces, many persons are already prophesying that we shall soon face a general shortage of labor.

Fortunately the labor supply is extremely flexible. As the need for labor grows, the number of workers offering their services in the market will also increase. Nothing approaching a *general* labor shortage is likely in the foreseeable future.

* A paper presented at the 103rd Annual Meeting of the American Statistical Association in joint session with the American Association for Labor Legislation, New York, December 29, 1941

Expanded war production has, however, created real shortages of workers of certain types and skills. These shortages are frequently acute in areas of rapid industrial growth. The specialized and concentrated demands of modern war have thus turned attention to ways of increasing the effective supply of labor in areas of defense concentration.

Broadly speaking, factories swamped with war orders can draw from four sources of labor supply:

First, from the existing supply of unemployed workers in the area. Even the most active defense centers have their unemployed—workers who, although clearly employable, cannot qualify for the available jobs because of age, sex, race or lack of needed skills. Many of these workers will eventually get jobs as the demand increases and hiring standards are further relaxed.

Second, through the transfer of workers presently engaged in other pursuits within the area. This process is stimulated by the higher earnings in war industries; reports indicate that considerable numbers of self-employed workers and workers in various service industries are already shifting to defense jobs. The transfer process is also hastened by employment reductions in non-defense industries resulting from the curtailment of essential materials—the so-called priorities unemployment.

Third, from non-worker groups among the resident population who are induced to enter the labor market by the prospect of easily-secured jobs at high wages. The number of persons who may decide to become workers during a war-time boom is very large—for the country as a whole it runs into the millions. Recent studies show that non-workers have already entered the market in large numbers in many centers of defense activity, and it may be expected that the declaration of war will accelerate this movement.

Fourth, from the supply of migrant workers attracted to the area by reports that jobs are plentiful and that good wages are being paid. In many areas the available and potential local labor supply appears ample to meet all prospective needs. In other areas, however—notably the sites of large construction jobs and a number of very rapidly expanding shipbuilding and industrial towns—the resident supply of labor is clearly inadequate.

The sharp concentration of defense contract awards is altering our previous industrial pattern, thus necessitating the migration of labor on a considerable scale. The accelerating impact of priorities unemployment, superimposed upon existing unemployment, is creating rather formidable labor surpluses in many areas, thus developing an expulsive force which adds to the total volume of movement.

The result has been a sharp stimulation of movement to many areas of the country. This defense migration has been brought vividly to public attention by the rapid massing of large numbers of construction workers in isolated army camp and powder plant towns.

The needs of war have thus transformed the migrant "bum" of the depression to the respected war worker of today. The border patrols and restrictive legislation of a few years ago are being replaced by a variety of devices to encourage migration, including advertising by private contractors, expansion of the public employment office clearance system, and the defense housing and community facility programs. Poor housing, overcrowding, health hazards, skyrocketing rents and inadequate school, sewer and water systems are now matters of public concern, partly because it is feared that such undesirable conditions may discourage the migration of needed workers.

It is clear that defense migration is raising increasingly serious problems. As in the case of depression migration, however, there is a general lack of knowledge concerning the development. Little definite information is available as to the extent of the movement, its success, or the types of persons moving. Once again imagination has been substituted for research. The wild tales of the thirties, which had hundreds of thousands of half-wild boys and girls roaming the country, are now replaced by equally exaggerated reports that all America is on the march.

In an endeavor to substitute fact for fiction, the Federal Security Agency some months ago requested that the WPA Division of Research undertake a series of studies of migration to defense areas. The surveys were designed primarily to determine how many workers and persons had moved to the area during the past year, where they had come from, what types of people they were, the occupations and industries in which they had been employed, the success of various groups of migrants in finding employment after migration and the extent to which they had shifted to new occupations and industries after migrating. The surveys covered the activities of civilian workers only, and no attempt was made to gather information about persons who had left the survey city during the year.

The information was secured through a sample census of each area, using techniques generally similar to those of the WPA Monthly Report of Unemployment.[1] Particular attention was given to coverage of rooming houses, lower-priced hotels, defense housing projects, and tourist and trailer camps.

In all, 51 areas were selected for survey. Most of the cities selected

[1] See "Dynamics of Labor Supply," by Howard B. Myers, this JOURNAL, Vol. 36, pp. 175–184, June, 1941.

had received large war material orders or construction contracts; a number of towns with few defense contracts were included as control areas. Data are now available for 40 cities and surveys are in process in the remaining 11 cities. The results of the survey are presented in a brief mimeographed report for each city.

The data thus far available appear to warrant a number of observations concerning defense migration, pre-war phase. It should be said at the outset that generalization concerning the movement is hazardous. The situation varies markedly by locality, depending on such factors as the type and intensity of defense activity, the size of the resident labor supply, the economic situation in nearby areas, and the ability of the community to house and service the incoming population.

Before going further, I should point out that by no means all migration to defense areas is defense migration in the narrow sense. Broadly speaking, one nondefense worker moves to a defense town for every worker who comes in to take defense employment. The enticing prospect of a job draws clerical and service workers as well as skilled metal tradesmen, operatives and construction workers. Many of these workers secure jobs which contribute indirectly to the war effort, others join the ranks of the resident unemployed. All contribute, however, to the social and economic problems which migration raises and, consequently all are included in the discussion which follows.

In general, defense migration has been of two main types: (1) the movement of construction workers to camp and new facility sites, many of which have been located in rural or small-town areas; (2) the movement of workers to war industry centers, for the most part the larger cities. The industrial movement has been less spectacular, but is of longer duration and socially and economically is much more important. I shall discuss primarily this latter type.

Perhaps the primary point to make concerning defense migration to date is that, by and large, it has been a good deal smaller than many of the more frenzied newspaper and magazine stories would have us believe. The movement to large construction jobs has been impressive, of course, but most of this has been temporary. It is true, further, that a few industrial towns have experienced a hectic mushroom growth. Migrants into San Diego, California, total 27 per cent of the 1940 population—the highest rate among the 40 areas for which figures are now available. Wichita, Kansas, a booming aircraft center, is second, with a 20 per cent migration rate. While the rate of migration is lower, some of the larger cities have experienced truly astounding in-movements. For example, more than 150,000 persons have moved into Los Angeles and its satellite towns since October 1940. Since the same date more than

50,000 persons have moved into Washington, D. C., and more than 40,000 to Seattle, Washington.

These cases are exceptional, however. In more than half of the 40 areas for which data are available the migration rate has been less than 5 per cent, and in only 9 of the 40 cities has it been 10 per cent or more.

On the other hand, the surveys indicate that the rate of migration has been *increasing* in most areas. It is probable that it will grow even more rapidly during the coming months, stimulated by the marked intensification of the war effort, by the near-absorption of the resident labor supply in certain "hot" areas, and by rapidly-growing priorities unemployment. During the first year of war, migration should exceed by a considerable margin the volume during the pre-war period.

Second, it is pleasing to report that defense migration thus far has been, on the whole, strikingly successful. In half of the areas surveyed the unemployment rate for all migrant workers is less than 8 per cent; in a fourth of the areas it is less than 5 per cent. Only one city out of seven has a migrant unemployment rate of 15 per cent or more. The highest rate among all cities recently surveyed is 17 per cent, for Fort Smith, Arkansas, where large numbers of workers had flocked in in anticipation of the start of work on a new army camp. Among the very large cities, St. Louis and Los Angeles are highest, each with an unemployment rate of 16 per cent.

In view of the almost completely unguided nature of the movement, and considering the fact that the surveys included considerable numbers of migrants who had very recently arrived in the area and had had little opportunity to adjust themselves, the unemployment rates reported are surprisingly low. Defense migration thus presents a welcome contrast to the tragic experiences of migrants during the depression.

Not only have the great majority of the defense migrants obtained jobs, large numbers of them have got better jobs than they held before migrating. Occupational upgrading has been widespread. Shifts among manual workers from unskilled to semi-skilled, and from semi-skilled to skilled have been especially frequent. As a result of this process the proportion of migrants working at unskilled occupations is surprisingly small—in the great majority of towns less than 10 per cent. Income data were not obtained, but in view of the occupational upgrading reported and the relatively high wages and full employment in most defense lines, it seems clear that the incomes of a large proportion of the migrants have risen.

Although migrants in general have been quite successful in finding jobs, certain groups have fared less well than others. Women have been far less successful than men in obtaining jobs—in most areas their rate

of unemployment is 3 or more times that for men. Negroes, too, have been relatively unsuccessful as migrants. Their unemployment rate is 3 or more times the rate for whites in most areas to which Negroes have migrated in appreciable numbers.

Young workers have been generally more successful than their elders. The very young group is a striking exception. In nearly every area workers under 20 reported the highest unemployment rate of any age group. Most of these youth have, of course, entered the labor market recently, and have had little or no previous work experience.

As would be expected the migrants who have been in the area longest tend to have the lowest unemployment rates. In general, migrants who have come from nearby areas tend to report less unemployment than those who have travelled far, probably because the former group more often return home if they fail to get a job.

In nearly all areas skilled manual workers and professional and technical workers have the lowest unemployment rates among the migrants. The least successful occupational groups are nearly always the service workers, particularly domestics, who often report extremely high rates of unemployment.

Negroes make up only a small proportion of the migrants to defense centers. In most areas fewer than 5 per cent of the migrants are Negroes, and even in the South, migration rates for Negroes are much lower than for whites. This is understandable, in view of the widespread discrimination against Negroes in war industries. It contrasts sharply with experience in the first World War, however, when a large scale migration of Negroes to Northern industrial centers took place. As the demand for labor increases and present employment restrictions are relaxed, it is probable that Negroes will begin to move in greater numbers.

Contrary to popular impression, relatively few of the migrants are coming from agriculture. This is rather surprising in view of our large agricultural labor reserve. In most defense centers, however, fewer than 10 per cent of the migrants are farm workers. Even in the South, the proportion is seldom as high as 15 per cent.

Defense centers thus far have secured their workers primarily from urban areas. Most of the rural migrants have come from villages; the proportion from the open country is very small.

Few of the migrants have travelled far; in most centers the average distance is less than 125 miles. The California cities are outstanding exceptions to the general rule. Migrants to Long Beach have moved an average of more than 1,000 miles, while Los Angeles migrants have

averaged nearly 1,300 miles. Workers from mid-Western cities make up a large proportion of these long distance migrants.

Migrants as a group, are young. In nearly half of the cities the average age of all migrant workers is less than 30 years; in no city does the average rise as high as 35 years. The figures reflect both the greater mobility of young workers and the low hiring age limits in many defense industries. In the aircraft town of Wichita, Kansas, where hiring restrictions are unusually severe, the average age of all migrant workers is under 25 years.

The migration surveys provide evidence that the rising demand for labor in defense centers is drawing non-workers into the labor market in considerable numbers. In most of the cities surveyed from 10 to 20 per cent of the migrant workers had never had a job at their previous residence. Most of these persons were students and housewives entering the labor market for the first time. The employment record of this group is generally poor; in most cities the proportion who have obtained jobs is definitely smaller than for migrants with work experience.

The proportion of one-person families among the migrants at their new locations is extremely high, ranging from 30 to 50 per cent for most areas and reaching a peak of 78 per cent for Washington, D. C. It is well known, of course, that single persons are highly mobile. Large numbers of these one-person families are incomplete, however; in most towns the proportion is between 20 and 50 per cent of all one-person families. In other words, from a fifth to a half of these workers left their families behind when they migrated. In part this separation reflects the normal instability of the migration process—the bread winner leaving his family behind until he settles in a new location. In part, however, it results from the serious housing shortages existing in many defense areas.

A striking relationship between the proportion of migrant families which are incomplete and the availability of housing is afforded by a comparison of the migration data with the findings of a series of residential vacancy surveys recently conducted in many of the same areas by the WPA Division of Research. As the residential vacancy rate declines the proportion of migrant families which are incomplete rises sharply. It is clear that many migrants leave their families behind simply because they are unable to find housing accommodations for them. Inadequate housing thus increases the proportion of migrants who are unstable and, in consequence, tends to increase the labor turnover rate in war industries.

The extent of "doubling up" among multi-person families provides a

further index of the congested housing conditions in defense areas. Doubling up is common in all of the defense areas surveyed; in most areas 30 per cent or more of the migrant multi-person families are sharing their dwelling with other persons.

While the extent of defense migration thus far has frequently been exaggerated, the WPA surveys show conclusively that considerable population movements have already taken place and that migration is still on the upswing. An all-out war effort, accompanied by rapid shifts in industrial concentration, vast new demands for labor in some areas and large scale priorities unemployment in other areas, is certain to provide a marked stimulus to the further movement of population.

We are clearly in for large scale migration of labor during the next few years. In many respects this migration may differ from that experienced to date. It appears probable, for example, that as the demand for labor grows migration from rural areas will increase, the movement of women and Negroes will also grow, and these groups will be more successful than heretofore in obtaining jobs.

Migration is already creating serious economic and social problems in many defense communities. It is important that these be solved and that they be solved quickly.

Meantime we should not forget the post-war migrant problems now being created. It seems probable that many war industries will shrink as rapidly as they have arisen, that the present extreme concentration of industrial activity will be to some extent reduced, and that the country will again face large-scale unemployment. These changes will provide new stimuli to migration. The course of post-war migration, however, will be far less happy than that of the present day. Post-war migration may well be depression migration all over again—the type with which we have become all too familiar during the decade just ended.

As the various expedients of the depression years abundantly illustrate, the wise direction of large-scale population movements is an extremely difficult task. To keep suffering to a minimum and to avoid the stagnation of pools of unused labor in new depressed areas requires careful planning over a considerable period of time. Whether such planning will be undertaken on a scale adequate to the problem remains to be seen.

ON THE SAMPLE SURVEY OF UNEMPLOYMENT*

By Lester R. Frankel and J. Stevens Stock
Work Projects Administration

In December 1939 the Division of Research of the Work Projects Administration put into operation a regular monthly survey of unemployment to provide accurate measurements of unemployment, employment and the labor force. Estimates are obtained directly through actual field counts of a cross-section sample of the total population each month. The superiority of this method rises from the fact that unemployment is measured directly rather than estimated indirectly as is necessary in the method used in other periodic unemployment estimates. No additions or subtractions, based upon assumptions as to unemployment reported in the 1930 or 1940 censuses, net average increase in the labor force, or changing age structure are necessary when direct measurement is made. The result of the sampling and direct interview procedure is an extremely fast and sensitive report on the state of the labor market as of one week each month. In some respects the procedures followed are similar to those of public opinion polls.

The Monthly Report of Unemployment operates in a national sample of 64 counties located in 45 states. Within each of these 64 counties a sample of households is selected for interview. These households are enumerated from four to six months in succession and then the sample households are changed so as not to arouse irritation through too prolonged inquiry.

The questions asked provide information on the size of the household, the number of persons under and over 14 years of age, and, for all persons 14 years and older, their relationship to the labor market during the census week, that is, the entire week immediately preceding the week of interview. Each person is asked one or more simple questions to determine whether he or she was employed, unemployed or out of the labor market. In addition to figures on the volume of unemployment and employment, data are secured on the age, sex, duration of unemployment for the unemployed, industry of the employed, and other characteristics of workers and non-workers in the population.

Beginning with the first day of the enumeration week, field schedules are tallied daily in each county. At the close of the enumeration week, a summary of the salient facts is mailed directly to the Washington office for consolidation and interpretation. These summary figures are released in memorandum form usually before the close of each month;

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 28, 1941.

further index of the congested housing conditions in defense areas. Doubling up is common in all of the defense areas surveyed; in most areas 30 per cent or more of the migrant multi-person families are sharing their dwelling with other persons.

While the extent of defense migration thus far has frequently been exaggerated, the WPA surveys show conclusively that considerable population movements have already taken place and that migration is still on the upswing. An all-out war effort, accompanied by rapid shifts in industrial concentration, vast new demands for labor in some areas and large scale priorities unemployment in other areas, is certain to provide a marked stimulus to the further movement of population.

We are clearly in for large scale migration of labor during the next few years. In many respects this migration may differ from that experienced to date. It appears probable, for example, that as the demand for labor grows migration from rural areas will increase, the movement of women and Negroes will also grow, and these groups will be more successful than heretofore in obtaining jobs.

Migration is already creating serious economic and social problems in many defense communities. It is important that these be solved and that they be solved quickly.

Meantime we should not forget the post-war migrant problems now being created. It seems probable that many war industries will shrink as rapidly as they have arisen, that the present extreme concentration of industrial activity will be to some extent reduced, and that the country will again face large-scale unemployment. These changes will provide new stimuli to migration. The course of post-war migration, however, will be far less happy than that of the present day. Post-war migration may well be depression migration all over again—the type with which we have become all too familiar during the decade just ended.

As the various expedients of the depression years abundantly illustrate, the wise direction of large-scale population movements is an extremely difficult task. To keep suffering to a minimum and to avoid the stagnation of pools of unused labor in new depressed areas requires careful planning over a considerable period of time. Whether such planning will be undertaken on a scale adequate to the problem remains to be seen.

ON THE SAMPLE SURVEY OF UNEMPLOYMENT*

By Lester R. Frankel and J. Stevens Stock
Work Projects Administration

In December 1939 the Division of Research of the Work Projects Administration put into operation a regular monthly survey of unemployment to provide accurate measurements of unemployment, employment and the labor force. Estimates are obtained directly through actual field counts of a cross-section sample of the total population each month. The superiority of this method rises from the fact that unemployment is measured directly rather than estimated indirectly as is necessary in the method used in other periodic unemployment estimates. No additions or subtractions, based upon assumptions as to unemployment reported in the 1930 or 1940 censuses, net average increase in the labor force, or changing age structure are necessary when direct measurement is made. The result of the sampling and direct interview procedure is an extremely fast and sensitive report on the state of the labor market as of one week each month. In some respects the procedures followed are similar to those of public opinion polls.

The Monthly Report of Unemployment operates in a national sample of 64 counties located in 45 states. Within each of these 64 counties a sample of households is selected for interview. These households are enumerated from four to six months in succession and then the sample households are changed so as not to arouse irritation through too prolonged inquiry.

The questions asked provide information on the size of the household, the number of persons under and over 14 years of age, and, for all persons 14 years and older, their relationship to the labor market during the census week, that is, the entire week immediately preceding the week of interview. Each person is asked one or more simple questions to determine whether he or she was employed, unemployed or out of the labor market. In addition to figures on the volume of unemployment and employment, data are secured on the age, sex, duration of unemployment for the unemployed, industry of the employed, and other characteristics of workers and non-workers in the population.

Beginning with the first day of the enumeration week, field schedules are tallied daily in each county. At the close of the enumeration week, a summary of the salient facts is mailed directly to the Washington office for consolidation and interpretation. These summary figures are released in memorandum form usually before the close of each month;

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 28, 1941.

an average of two weeks elapses between the first interview in the field and the completed memorandum.

After completing the preliminary field tallies, each county supervisor mails the completed schedules to a central machine tabulating office where additional information is taken from the schedules. These additional facts provide the basis for a much more detailed analysis.

In designing the sample, administrative conditions as to time, cost and centralized technical control were kept in mind. To meet these limitations the methods of stratified sampling and subsampling were adopted, involving selection of a relatively small number of primary sampling units (counties) and a subsample of secondary sampling units (households) within each of the larger units. To insure maximum statistical efficiency, both the primary and secondary sampling units were selected on the basis of a stratified sample.

The county was chosen as the primary sampling unit because of the following considerations: The county was a small self-contained geographic unit for which it was possible to obtain statistical information that could be used for stratification purposes. The variance of the unemployment rate between counties was low. Excellent county maps were available that made the county a very satisfactory unit from which to subsample. Finally, administrative operations in the average size county could be handled easily by a local supervisor with six or seven enumerators.

Both administrative and statistical conditions were taken account of in the selection of the secondary sampling unit. The household was chosen mainly because it was the smallest unit that could be sampled readily within a county. Such units were easily defined and located, whereas it would have been practically impossible to select a random sample of individuals without first having listed all persons within the county. The household was desirable also because of certain statistical reasons. A great deal of valuable household data, such as size, number of workers, and employment status, could be obtained only by enumerating household units. Moreover, for a great many characteristics, the household is a more efficient sampling unit than the individual because of the heterogeneity of individuals within households.

All of the counties in the United States were stratified by (1) size, (2) location, and (3) economic situation. Because of the great difference in county characteristics associated with population size, counties were first divided into the following strata:

(1) The nine counties in which the five largest cities are located (containing 14 per cent of the United States population).

(2) The 447 counties with 1930 populations of 45,000 and over (con-

taining 50 per cent of the United States population). These counties are referred to herein as "urban" counties.

(3) The 2,641 counties with 1930 populations of less than 45,000 population (containing 30 per cent of the United States population). These counties are referred to herein as "rural" counties.

Because of the great heterogeneity among the five largest cities, no attempt was made to select a sample of these nine counties; all were included in the sample. Samples of counties were selected from each of the other two strata. The method of selection was the same for both the "urban" and "rural" counties and, therefore, the procedure will be explained in terms of the urban counties only.

All urban counties were further classified by population size into three substrata (terciles): small, medium, and large. The same counties were then cross-stratified by states into three broad geographic locations: Northeast, Southeast, and West, so that there was about an equal number of counties within each of these locations.

The percentage of persons unemployed in each of these counties according to the 1937 Census of Unemployment was used as the third mode of stratification. At that time these data were the most current and reliable which were available for all of the counties in the United States. The counties were then divided into three strata (terciles) according to economic situation.

The urban counties were stratified into 27 cells: three population size groups, by three geographic locations, and by three economic levels. One county was selected from each of 26 of these cells, and two counties selected from the 27th cell because of the relatively high proportion of total urban population in this cell. Within the 27 cells, the selection of counties was at random except that a deliberate effort was made to maximize state coverage.

Twenty-seven counties were selected from the rural counties in a similar manner, using the same modes of stratification.

In all, 64 counties were selected for the sample—9 metropolitan counties, 28 urban counties, and 27 rural counties.

The household sample within counties was stratified according to population distribution. In each sample county the number of households selected within each city, town, village, or other minor civil division was proportionate to the population in those places.

The procedure for selecting the household sample was dependent upon whether the unit fell in an urban place (city, town, or village) or in the open country. In the former, a sample of city blocks was chosen by numbering a map and selecting every nth block; all dwelling units in these blocks were then listed and a subsample of households was

selected from these block lists. The sampling ratio was fixed so that there was an average of no more than 3 households selected per block.

In open country areas the household sample was allocated among the townships or similar minor civil divisions in multiples of 5. In each township one mile sections were chosen at random in the Washington office from maps of the sample counties. Five households are interviewed in each of these sections. The order of selection of households was determined arbitrarily by starting at the southwest corner of the section and proceeding around the section in a counterclockwise direction. Alternate sections were indicated for each township to be used when the original sections did not yield five households. The sections were numbered consecutively at random over the entire county and the enumerator proceeds from the lowest numbered section to the next higher in a township, interviewing up to 5 households per section until the quota for the township is obtained.

The intracounty sample is changed every four to six months. This is a compromise between having a constant group of households to study monthly changes, and having new households to keep the sample dynamic and to overcome interviewee resistance. It excludes persons in institutions and the armed forces.

National estimates of labor force, employment, and unemployment are made by expressing these characteristics in each sample county as a percentage of the total population interviewed, and then proceeding to a weighted average of these percentages for the country as a whole. The weights are proportionate to the populations of the cell or segment of the universe each county is chosen to represent. Thus, the weight of each of the nine metropolitan counties is proportionate to the total population of each of those counties. The weight applied to percentages derived from a county in any urban or rural cell is proportionate to the total population of all counties in that particular cell.[1]

Thus, national estimates are obtained directly in terms of percentages of total population. The total population in this case is defined as all persons not in the armed forces or in institutions.

In presenting labor market data from this survey, the results in percentage form are converted from total population to the population 14 years of age and over. Estimates of national totals are presented in monthly memoranda in absolute as well as percentage form.

[1] By the modes and levels of stratification outlined above, it is possible to obtain national estimates with as few as 27 counties. This may be done by selecting the 9 metropolitan counties, 9 urban counties and 9 rural counties. Both the urban and rural counties must be selected in each case in the form of a Latin square. From information obtained from the 9 urban counties, for example, it is possible to provide estimates for each of the 27 urban cells. Estimates for the urban and rural cells, and metropolitan counties would then be averaged together in the manner described above.

2.41.
18.1.6

PROPOSED ANNUAL SAMPLE CENSUS OF POPULATION*

BY PHILIP M. HAUSER
Bureau of the Census

THE IMPACT of World War II on this Nation, necessitating in rapid succession the national defense, lend-lease, and victory programs, has created an urgent need for social and economic statistics. It is fortunate that the Sixteenth Decennial Census was taken and is being completed during this period of national emergency. As a result we are able to mobilize our human and material resources on a sounder foundation of facts than has ever been possible before in a similar period.

Great economic and social changes, evidenced by large population shifts and industrial and occupational dislocations, have occurred, however, with a suddenness which defies the usual statistical description. The nearly two years which have elapsed since April 1, 1940 (the date of the census), are unparalleled in our history with respect to the velocity and extent of change which has occurred. In many communities the changes that have taken place in the number and composition of the population, in the employment status and industrial and occupational composition of the labor force, and in the use of and demand for housing have been greater since April 1, 1940, than in the entire preceding decade. As a result, there is an urgent need for current data, tied to the Sixteenth Decennial Census base, for administrative, legislative, industrial, and scientific purposes.

The Need for Current Statistics. The Bureau of the Census has, logically enough, been the center of a deluge of inquiries for current economic and social statistics from national defense agencies, regular government agencies, industry, business, scientific bodies, and the public at large.

For example, there has been considerable demand for postcensal population estimates for the states and cities—estimates which cannot be made through the usual techniques because of the volume and character of recent internal migration. Such statistics have a wide variety of uses, manifested in the requests for the data which have been received. The Vital Statistics Division of the Bureau of the Census itself, as well as the United States Public Health Service and state and local public health offices, are in need of population figures as a basis for computing current fertility, mortality, and morbidity rates. State and local governments and planning commissions require current population figures for

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 29, 1941.

the efficient administration and planning of public services. Postcensal population figures are needed by a number of agencies as the basis for the allocation of funds, as, for example, the allocation of funds by the Social Security Board to the states for public health work. Finally, numerous requests also have been received for current population statistics from industrial and business concerns and advertising agencies which are faced with the problem of making adjustments not only to new production needs and priority regulations, but also to rapidly changing markets.

An even more pressing demand for postcensal population figures is found among the defense and war agencies. The large and unknown population shifts affect plans of the War and Navy Departments, the Office of Civilian Defense, and the War Production Board. The aircraft carrier and the airplane have made large sections of our coasts vulnerable to attack, and the adequate protection and mobilization of these vital target areas are facilitated by a knowledge of their population aggregations and composition.

There is need also for current data relating to the labor force. The 1940 population statistics relating to labor were designed largely to measure the volume of unemployment and the characteristics of the unemployed. At the present time there is urgent demand for figures relating to the potential labor force—that part of our population not normally engaged in productive work for pay or profit which must be drawn into industry to meet the demands of our war economy. Moreover, there is need for current information on the occupational and industrial affiliation of workers for use in connection with the planning of production, the training of workers in occupations where shortages exist, the transfer of workers into new productive channels with maximum efficiency and a minimum of dislocation, and the location of new plants.

Great movements of population to important production or administrative centers have created new demands for housing requiring great outlays for new construction. The housing agencies in Washington and private construction organizations have made urgent requests for current housing statistics as a basis for planning new housing projects in a manner to meet the needs of the present emergency with a minimum of long-run difficulties. Up-to-date census statistics relating to rent would not only provide direct information on the changing cost of living, but would also be important as a base for the continuation of current indexes of rent changes.

Finally, there is a demand for current statistics from various agencies, including the National Resources Planning Board, to serve as a basis

for post-war readjustment. It may be safely stated that in the post-war readjustment period the need for statistics as a basis for action, administrative and legislative, will be as acute as the need for data in the emergency war situation. It will be necessary to have not only the basic statistics of the 1940 Census to tell us where we were at the onset of the war, but also current statistics to show where we have gone. It is only with these combined types of facts that adequate planning and administration can be effected during the trying times ahead.

These few illustrations of the demand for current population statistics suffice, perhaps, to make clear the need for such data for our immediate national defense and war efforts and for post-war readjustment purposes, as well as for more "normal" important governmental, industrial, and scientific uses.

Recognizing the importance of current statistics as a basis for mobilizing our human and material resources in the war emergency, and to meet the varied urgent demand for such data, the Bureau of the Census, in its role as a service agency, has devised a plan for the conduct of a sample census of population (and also of agriculture). Such a sample census might be taken on an annual basis, at least during the period of rapid change which confronts us. This plan is as yet entirely unofficial and has the character of a staff plan which could quickly be put into operation if desired. The essentials of this plan are briefly outlined in the paragraphs which follow.

The Type of Data to Be Obtained. The first schedule to be used if a sample census of population were taken would probably include inquiries which could be classified under three main headings: (1) General population, (2) the labor force, and (3) housing.

The questions relating to population would include age, sex, color, marital status, relationship to head, and place of residence on April 1, 1940, the date of the last census. The last of these questions would be designed to obtain the volume of in- and out-migration by comparing place of residence at the time of the sample census with place of residence at the decennial census.

The inquiries relating to the labor force might, in the first sample census, include employment status questions, similar to those on the 1940 population schedule, designed to indicate the number of persons in the labor force by their work status. These questions would also result in statistics relating to persons not in the labor force, showing their status during the cross-section week studied, and would be supplemented by inquiries designed to provide information indicating the potentialities of persons not in the labor force for labor force activity. The labor force inquiries might also include questions relating to occu-

pation, industry, and class of worker of persons in the labor force, with provision to obtain the occupation of persons not in the labor force.

The inquiries relating to housing would, in addition to the usual information indicating location, include questions on type of structure, general condition of the structure, age of structure, occupancy status of dwelling unit, number of rooms, and monthly rent for the tenant-occupied units, and estimated monthly rent for owner-occupied units.

These are the types of inquiries which it would be feasible to include in the proposed sample census. The final schedule would, of course, be determined only after consultation and clearance with consumers of census data. If the sample census were taken on an annual basis, a number of questions could be rotated from year to year so that the scope of the inquiries would be greatly expanded and proper interrelations of data obtained.

Estimates of Total Population. The most difficult question which the sample census would be called upon to answer is probably that relating to the total population of the area studied. Normally, estimates of the total population of an area would be dependent upon the use of an inflation factor obtained by comparing the population of the area studied on the sample census date with its population on the decennial census date. This method, however, would not only be expensive (because of the problem of obtaining decennial census figures for the areal units in the sample), but would, because of the relatively slow rate of change in total population, result in large sampling errors.

To meet this problem, it is planned to estimate the total population of a place by the "migration method." This would involve the application to the sample census data of the same technique designed for making intercensal population estimates from the 1940 Population Census migration statistics. The method essentially involves estimating the population of a place by adding to the population of the place as reported in the last census the number of births and in-migrants, and subtracting the number of deaths and out-migrants.

It is obvious that through the use of this method the error of the total population estimate is confined to the error in estimating migration and in the reporting of births and deaths. Approximate error formulas have been developed which, together with empirical studies which have been conducted, indicate that the migration method of estimating total population will usually be more precise than the method dependent on knowing the population of the areal units in the sample and in the preceding census.[1]

[1] This work has been done mainly by Morris H. Hansen, Acting Assistant Chief Statistician, Division of Statistical Research, Bureau of the Census.

It is planned to make statistics available for the United States, its regions and geographic divisions, the states, and cities having 100,000 or more inhabitants, or possibly 250,000 or more inhabitants, depending on the funds made available. To show areas smaller than these will require relatively large additional outlays probably beyond the realm of possibility for a reasonable sample census undertaking. The details with which the statistics will be reported for these various areas will, of course, vary with the size of the population, with the greatest amount of cross-classification being made available only for the United States as a whole or large subdivisions thereof.

The Design of the Sample.[2] The basic problem in the design of the sample for the proposed sample census of population is, of course, the determination of the character and number of sampling units to be used. Ideally, it would be desirable to select the same kind of sample as was selected during the 1940 Decennial Census of Population.[3] This would, of course, be entirely impracticable because of the huge expenditure which would be involved in obtaining the basic listing of the population.

The only feasible method of sampling, therefore, lies in the selection of geographic units of some kind. In selecting the sample areas the usual dilemma must be faced. Small areas are, of course, more efficient from a sampling standpoint, but large areas are easier to handle and cheaper to enumerate and tabulate. In designing the actual sampling plan contemplated, a compromise plan was effected which was calculated to have the maximum sampling efficiency within the limits of a reasonable expenditure of time and funds.

The proposed sample will consist of small areas scattered throughout the country and containing approximately 100 persons each. These areas, termed "parcels," will, in cities of 50,000 or more inhabitants, consist of city blocks or pieces of blocks. In the remainder of the country the parcels will consist of segments of the enumeration districts that were used in the 1940 Census. In rural parts of the country that have been surveyed and laid out on a township and range basis, the parcels may consist of sections or groups of sections.

The method of selecting the sample may be briefly outlined as follows: In cities of 50,000 or over, the large blocks (which might be defined as those containing 40 dwelling units or over) will be divided into two or more parcels. The sample parcels will then consist of whole small blocks, and pieces of large blocks, drawn by lot. The drawing

[2] The design of the sample is largely the work of Morris H. Hansen and W. Edwards Deming.

[3] F. F. Stephan, W. Edwards Deming and Morris H. Hansen, "The Sampling Procedure of the 1940 Population Census," this JOURNAL, December 1940.

will take place in such a way that the sample parcels are stratified geographically, which automatically gives fair representation with regard to a number of other factors.

In the territory outside the cities of 50,000 or over, where the blocks are not numbered, the enumeration districts will be stratified and divided into parcels as near the required size as is possible in the time allotted for preparation. The parcels for the sample will then be drawn by lot, proportionately from each stratum. The stratification will be geographic, and by its characterization as urban or rural, farm or nonfarm, in or not in a metropolitan district, by rent, color, and possibly by other characteristics. Although parcels so selected would constitute the main bulk of the sample, it would be necessary to draw samples from certain special universes, such as large institutions, prisons, army posts, large new apartment houses, etc., to guard against large sampling errors and administrative difficulties. Techniques have been devised to sample these special universes as may be required in different areas.

Precision of the Sample. On the basis of theory and actual experimentation, it is now possible to predict in advance the precision of the sample for the various characteristics and areas. Some important research[1] has been conducted in the Bureau of the Census in the determination of the most efficient sampling units. Studies have been made on the effect of stratification and size of parcel on the average intraclass correlation for various population characteristics. Knowledge of the behavior of the intraclass correlation coefficients enables sampling errors to be predicted for various sizes and kinds of sampling units, various sizes of sample, and various population characteristics. The work which has been done comparing predicted errors with observed errors in experimental sampling studies has demonstrated that the intraclass correlation coefficients have sufficient temporal and geographic stability to permit useful estimation of the sample errors.

In general, it may be stated that the results to be expected from the proposed sample, whether taken on an average 2 per cent or 5 per cent coverage (with variable sampling ratios for areas of different size), will have a relatively small sampling error and be adequate to serve as a practical basis for action for many administrative, legislative, or scientific purposes.

For example, it can be demonstrated that for a city having one million or more inhabitants with a normal amount of migration, the total population based on a 5 per cent sample can, on the average, be estimated with an error of less than four-tenths per cent and with an error

[1] This research has been conducted largely by Morris H. Hansen and William N. Hurwitz. (See pp. 89–94 *infra*.)

that will very rarely be as great as 1 or 2 per cent. Populations of smaller cities can be estimated with proportionately greater error.

Sampling errors will be smaller in estimating percentages than in estimating absolute amounts. This follows because the estimate of the absolute amounts will involve the use of an inflation factor which is also dependent on an estimate (that of the estimated total population of the area). For many purposes the relationships as expressed in percentages are the important desiderata, and, from a 5 per cent sample for a place having 250,000 or more inhabitants, such relationships as can be expressed by percentages can be estimated on the whole with an average error of less than 1 per cent for general population and labor force characteristics.

The average relative error in estimating an absolute amount for population and labor force characteristics (other than total population) would, for a place of approximately 250,000 persons, be less than 10 per cent, assuming that the total population is estimated by the migration method.

Relation of Sample Census to Decennial Census. The proposed sample census would, of course, have a number of limitations, most important of which would be its failure to provide even total population figures or simple counts of characteristics for small areas. The need for a complete inventory of the Nation's population can be met only by a complete census, and in a period such as this it may be necessary to take a quinquennial census of population to obtain the needed data for local inventory purposes.

The sample census would have an important bearing, however, on the complete census, whether taken decennially or quinquennially, in that it would provide a better basis for planning and conducting it. The complete census could be better interpreted in the light of the knowledge of changes recorded from year to year in an annual sample census, while these changes are taking place. Moreover, the sample and complete censuses combined would provide a time series, the significance of which, of course, would transcend that of the sum of the individual points in the series.

Perhaps an adequate program for the decade which faces us is one in which a sample census of population is taken in 1942 and 1943, a quinquennial complete census in 1945, and additional sample censuses in 1947 and 1948. This spacing would permit sufficient time for the adequate planning and conduct of the complete censuses while providing a reasonably current series of statistics.

It is not without significance that the German war effort was preceded by a great expansion of statistical activity to serve as a basis for

the mobilization of the German Nation in its total war effort. This Nation might well profit from the German example in planning its own emergency statistical activities. Failure to provide reasonably current statistics on budgetary grounds may prove to be a shortsighted and expensive economy.

This brief statement outlines the essential features of the proposed sample census plan. The urgent demand for current statistics calls for the conduct of the first sample census in the near future if the statistical needs of the various war and defense agencies are to be served adequately. The plan is ready and can be put into effect whenever authorization and funds are made available. There is no doubt that the proposed sample census would provide current statistics of great use for the better planning and administration of our war effort (and post-war readjustment), and that would certainly be its immediate and major purpose.

RELATIVE EFFICIENCIES OF VARIOUS SAMPLING UNITS IN POPULATION INQUIRIES*

BY MORRIS H. HANSEN AND WILLIAM N. HURWITZ
Bureau of the Census

IN THE PLANNING of an annual sample census of population, agriculture, and housing, it has been necessary for the Census to consider the relative efficiency of different sizes of sampling units for estimating the characteristics of a finite population. We shall present some preliminary results that have been obtained in a few selected areas, and that are now being extended in a Work Projects Administration project sponsored by the Bureau of the Census. Later we shall present a summary of the work being done in this project.

The units of sampling that are being considered are: An individual person or household, or a small geographic area such as a city block, a segment of a block, a group of blocks, or a small rural area. Such sampling units will be referred to as clusters of elements, that is, clusters of dwelling units, farms, or people. For example, where a block is the sampling unit, the cluster of elements consists of all the persons or all the dwelling units in a block.

One of the principal problems which concerns us is to determine for a stratum which of these sampling units is most efficient, subject to cost and administrative limitations. We shall discuss the relative efficiency of different sampling units only in terms of the relative magnitudes of their sampling variances, where the variances are computed on the same unit basis.

When clusters of elements are the sampling units, the correlation between the elements within clusters will influence the sampling error. This can be seen from the formula[1] for the variance of a sample proportion or average, which involves the correlation between elements within

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 28, 1941.

[1] In the formula given, the number of elements is assumed to be the same for all clusters. Actually sampling units will vary in size, and the variance of a sample estimate must take account of the variability in the size of the cluster. If the average size of the cluster is $\bar{N}$, then a first approximation to the variance when clusters are the sampling units is

$$\frac{M-m}{M-1}\cdot\frac{PQ}{m\bar{N}}\left[1+\frac{\Sigma\rho_i N_i(N_i-1)}{M\bar{N}}\right] \text{ where } \rho_i=\frac{(P_i-P)^2-\dfrac{P_iQ_i}{N_i-1}}{PQ}.$$

The error in this approximation is given exactly by the expected value of

$$\left(\frac{\bar{N}_a^2}{\bar{N}^2}-1\right)\left(p_a-P\right)^2,$$

where $\bar{N}_a$ represents the average of the N_i for a sample of m clusters, and p_a is the corresponding sample proportion.

sampling units. The variance of a sample proportion is given by the formula:

$$\sigma_p^2 = \frac{M-m}{M-1}\,\frac{PQ}{mN}\,[1+\rho(N-1)]$$

where

N = number of elements per cluster,

P = proportion of elements in the stratum having a particular property,

M = total number of sampling units in the entire stratum,

m = number of clusters sampled,

p = proportion of elements in the sample having a particular property,

ρ = intra-class correlation between elements in the sampling unit,
$= E(x_{ij}-P)(x_{ik}-P)/E(x_{ij}-P)^2$, where the symbol E preceding a quantity denotes the expected value. When evaluated for the finite population ρ is equal to

$$\left[\frac{\sum (P_i-P)^2}{M} - \frac{\sum P_iQ_i}{M(N-1)}\right]\Big/PQ.$$

The summations are over all sampling units in the stratum.

The first factor, $(M-m)/(M-1)$, in the above formula for the variance is the finite population correction. The second factor PQ/mN, is the variance of a sample proportion if mN elements are sampled independently at random, with replacement. The third term is the factor which measures the contribution to the sampling variance of cluster sampling. If $N=1$, this third term becomes unity, and we have left the regular formula for sampling individuals at random. If N is greater than 1 (and M is large), we have $\rho(N-1)$ as the relative change in the sample variance from the variance obtained by sampling individuals at random. The influence of relatively small values of ρ, when N is large, is very striking. For example, if N is 100 and the correlation is only .05, the variance of a sample proportion is six times the variance obtained by sampling individuals. The relative efficiency in this case is one-sixth. If N is 1,000 and the correlation is .05, the variance is fifty times the variance obtained by sampling individuals.

For most population and housing items the correlation decreases as the size of the sample cluster increases. But usually the decrease in ρ is at a less rapid rate than the increase in N, so that, ordinarily, increases in the size of cluster lead to substantial reductions in efficiency. This will be illustrated later.

While usually the intra-class correlations encountered in practice are positive, and therefore the sampling of a cluster of elements is less

efficient than the sampling of individual elements, many important exceptions to this statement may be found.

Let us describe a physical situation which brings about a negative correlation, and therefore greater relative efficiency for a sampling unit containing a cluster of elements than for a single element. Assume that we wish to estimate the proportion of males in a population. First, let us restrict ourselves to 4-person households, each consisting of a husband and wife and two children. We shall assume that there are as many households in which both children are male as there are in which both children are female, and that these constitute half of the total households. The other half of the total households contain one child of each sex. This assumption agrees fairly well with the actual composition of 4-person households. The intra-household correlation for this assumed population is $-\frac{1}{6}$ or approximately $-.17$. This negative correlation reduces the variance when households are the sampling units to one-half the variance obtained when individuals are the sampling units, and hence the household is approximately twice as efficient as a sampling unit of a single person.

Physical situations for which the intra-class correlation is negative can be generalized as follows: If it is observed that an element in a cluster has a certain property, then it is less likely that the next element observed will have that property if it came from the same cluster than if it came from any other cluster.

Of course, in any actual population, households are not of the same size and do not follow exactly the sex distribution given in our hypothetical example. Results from actual observations, however, do not differ much from the results of our illustration. As examples, results are presented for two counties, one in Iowa and one in New Hampshire, each stratified by rural and urban within the county. In both the rural and urban parts of these counties we found that for estimating the sex distribution, the household sampling unit was approximately twice as efficient as the sampling unit of one person.

There are other characteristics for which the correlation between members of a household usually will be negative and for which the household will be a more efficient sampling unit than the individual. A similar situation, for instance, holds for the estimation of the proportion of persons 14 years old and over that are in the labor market. For example, in the county of Iowa considered above the relative efficiency for this characteristic was 1.4 in the urban part and 2.2 in the rural, and the corresponding values for the New Hampshire county were 1.0 and 1.5.

For practical purposes a sampling unit as small as a household may

not be administratively feasible. Moreover, the efficiency of sampling units for estimating the proportion of males in the population, or the proportion of persons in the labor market, is not typical of most population and housing characteristics. It is of practical importance to investigate efficiency of sampling units for other statistics and for clusters of elements larger than a household.

Results of analyses indicate that for most population and housing items a large size sampling unit is considerably less efficient than a small one. Such results were obtained, for example, from an analysis of 1930 data from Wheeling and Philadelphia for samples stratified by census tracts. This analysis shows that, for estimating the proportion of persons in specified age groups, say ages 5–14, blocks averaging between 150 and 200 persons, as sampling units, would be $\frac{1}{3}$ to $\frac{1}{2}$ as efficient as a sampling unit of one person.

For estimating the proportion of dwelling units that are owner-occupied, the block as a sampling unit in these two cities was found to be $\frac{1}{8}$ to $\frac{1}{3}$ as efficient as a single dwelling unit. This loss of efficiency arises from the fact that within a stratum most of the dwelling units in some blocks may be owner-occupied, while most of the dwelling units in other blocks within that same stratum may be tenant-occupied or vacant.

The efficiency of the block is quite different for estimating moderate vacancy rates. For example, in the same cities the block was found to be nearly as efficient as a household for estimating strata vacancy rates less than .02. Its efficiency decreases to approximately 3/10 to 7/10 for estimating vacancy rates over .02. These results arise from the fact that, at least for these cities, in 1939, the intra-block correlation for moderate vacancy rates is close to 0 while for vacancy rates above, say, .1, the correlation is relatively high.

In rural areas and small cities, segments of enumeration districts may serve as the sampling units. For estimating the proportion of males 5–14, in rural areas within counties, a cluster consisting of five households, or 20 persons, was found to be .8 as efficient as one person, clusters of 10 households slightly less, while an entire enumeration district averaging 250 households was found to be only 1/20 to 1/10 as efficient as one person.

The loss of efficiency with an increase in the size of sampling unit is particularly pronounced in estimating the proportion of dwelling units that are owner-occupied. The efficiency of a segment containing five households was about 8/10 that of a single household, and dropped to 1/10 when the enumeration district was considered as the sampling unit, and in some instances as low as 1/100.

It is to be pointed out that the relative efficiencies of different sampling units are not independent of the system of stratification.

The above results are tentative, being based on only small scale experiments, but they give a general idea as to the relative efficiencies of different sizes of sampling units for estimating population and housing items. They are now being extended in a WPA project which the Bureau of the Census is sponsoring in New York City. The study is based on more than 1,000,000 names selected from the 1930 Census schedules, and approximately 500,000 names from the 1940 schedules drawn from cities and counties throughout the United States.

The investigations are being extended not only to alternative random sampling designs but also to systematic designs such as taking every n-th person, or household, or block within a stratum. Thus, detailed information will become available concerning factors which influence the choice of a sampling design, such as gains from stratification, the utility of sub-sampling, and the effect of variability in the sizes of the sampling units. The results of the analysis of 1930 and 1940 data and some less extensive results for intra-census years will make it possible to determine the relative temporal and geographical stability of some of the factors which are involved in the sampling theory. Thus, the study of the relative efficiencies of different sizes of sampling units is only one phase of this WPA study.

We shall conclude with a word about the choice of a mathematical model. The model we are now using is a finite population, and the variances and correlations are evaluated for the finite population. There are other models, that might have been used, among which the one posed by Cochran, and by Yates and Zacopanay, is outstanding. As commonly used, the latter model assumes that the finite population is a sample from an infinite population; and if the population is divided into clusters, each element can be expressed as the sum of two quantities, one of which is constant for all elements in a cluster but varies independently from cluster to cluster, the other varies independently within clusters. This model, however, does not permit negative correlation, and hence a cluster of elements according to this formulation can never be more efficient than a single individual. This is not identical with some of the physical situations given above.

Suppose that the finite population within which we are sampling consists of M sampling units, each unit consisting of N elements. Suppose further than m such clusters are sampled. Then this model leads to

$$\sigma_n^2 = \frac{(M - m)}{M} \left\{ \frac{\sigma_b^2}{m} + \frac{\sigma_w^2}{mN} \right\}$$

as the variance of the sample proportion, p, around the finite population proportion; where σ_b^2 is the variance of the quantity which is constant for all the elements within a cluster but varies from cluster to cluster, and σ_u^2 is the variance of the quantity which varies within clusters.

Now suppose that mN elements are sampled individually. Then the variance of the sample proportion, p', around the finite population proportion is

$$\sigma_p'^2 = \frac{(M-m)}{Mm}\frac{\sigma^2}{N} \quad \text{where} \quad \sigma^2 = \frac{N(M-1)}{MN-1}\sigma_b^2 + \sigma_u^2.$$

It can be shown that $\sigma_p^2/\sigma_p'^2$ cannot be less than 1, and therefore, under the analysis of variance formulation, a sampling unit consisting of more than 1 element can never be more efficient than a sampling unit of a single element. This is in contradiction, as illustrated earlier, to the physical situation where the intra-class correlation between elements within a sampling unit is negative. A modification in the model, such as the one Yates and Zacopanay made to take care of the so-called "Competition" between elements within a sampling unit, would be necessary in order to have a model which more closely corresponds with all possible physical situations.

RECENT DEVELOPMENTS IN SAMPLING FOR AGRICULTURAL STATISTICS*

By George W. Snedecor, *Iowa State College*, and
Arnold J. King, *U. S. Department of Agriculture*

Our title is broader than the subject matter we propose to discuss. We shall limit our remarks to recent developments in the research conducted at Ames, Iowa, by a group of representatives of the Agricultural Marketing Service and Iowa State College.

Investigations in sampling are divided broadly into two categories: those pertaining to somewhat extensive surveys and those of the more restricted kind associated with experimentation. Although we have not recently been active in the study of experimental design, Cochran, Autrey and Cannon[1] have given an account of a design for a switch-over type of experiment in which three or four treatments are administered successively to each experimental animal. This plan makes it possible to estimate carry over effects and to adjust for them, thus providing accurate comparisons of the results of the treatments and unbiased estimates of the experimental errors. The design is especially valuable in experiments where the milk production of dairy cows is the variate investigated, because due allowance is made for decreasing flow with advance of the lactation period. Since three rations were to be fed to eighteen cows, each for six weeks, the cows were divided into six groups of three so that the members of a group were about equal in known performance. The three cows of each group were given the three rations in sequences forming a Latin square. The use of a pair of complementary 3×3 squares, three groups being assigned to each, provided the adjustment for carry over effects. The solution is effected by applying the principle of least squares to equations in which the observed milk yield of any cow is expressed as a linear function of the effects of the cow's yielding ability, the period, the current and previous rations, and the experimental error. The design attained a satisfactory precision in this feeding experiment, the standard error being 1.5 per cent of the mean.

Turning now to the more extensive types of sampling in agricultural statistics, we consider first the work which Messrs. Jessen, McCarty and Houseman have been doing, in which they have investigated the impact of cost on sampling design. The maximization of information for a set total expenditure may lead to notably different designs from those set up from no other than statistical considerations. The procedt-

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 28, 1941.

[1] "A Double Change-over Design for Dairy Cattle Feeding Experiments," by W. G. Cochran, K. M. Autrey, and C. Y. Cannon, *Journal of Dairy Science* 24: 937, November 1, 1941.

ure which has been used is to derive not only a variance function but also a cost function, then minimize the former for fixed values of the latter. Two examples can be given.

Jessen[2] has examined the efficiency of a sample of some 800 Iowa farms, the variates being numerous items of agro-economic interest such as acreage, production, income and expenditure. The sampling unit in this survey consisted of the farms for which farmsteads lay in a quarter section, this unit lending itself conveniently both to random selection and to expansion of the sample means into state totals. The time consumed by each questionnaire was about an hour. From examination of the variance function alone, it was evident that in this survey the quarter section sampling unit was the most efficient of the grids, but that the individual farm would be advantageous as a unit if the difficulties of sampling, expansion, etc., could be avoided.

The consideration of cost revealed that with smaller total expenditure and shorter questionnaires larger sampling units would be more efficient. Let us consider each of these conditions. High costs of travel require larger sampling units. Although this increased taking of schedules in each chosen locality is statistically less efficient than the more diffuse placement of smaller units, it is economical of mileage. Hence, it is efficient to sacrifice some information for each schedule in order to obtain a greater number of returns cheaply. Similarly, short questionnaires make for larger sampling units, because, the cost per schedule being small, it is possible to secure many in every location. The increase in the number of schedules thus obtained compensates somewhat for the lower statistical efficiency per schedule. More restricted funds for sampling likewise force the taking of larger as well as of fewer sampling units. Thus, information per unit, as well as total information, is sacrificed to cost.

This finding emphasizes the desirability of small sampling units for the long questionnaires usually required in surveys such as those taken in farm management research. Here the individual farm is usually indicated as the most efficient unit, and may also be the most practical one if ways can be found to make effective random choice. This is especially true if expansion of the sample means is not desired.

A second example of a sample cost study is reported by King, McCarty and McPeek.[3] This study is based on the 1940 pre-harvest wheat survey in Kansas, a route sampling in which the sampling unit is

[2] "Statistical Investigation of a Sample Survey for Obtaining Agro-economic Facts," by Raymond J. Jessen accepted for publication as a research bulletin by the Iowa Agricultural Experiment Station.

[3] "An Objective Method of Sampling Wheat Fields to Estimate Production and Quality of Wheat," by Arnold J. King, Dale E. McCarty, and Miles McPeek, accepted for publication as a U.S.D.A. Technical Bulletin.

designated as one field to each mile of wheat along a checkerboard route. In every such field two patches of wheat, each of about one five-thousandth of an acre in area, and selected at random, were harvested. The Kansas sample comprised 1,100 fields.

The variance function set up by Mr. E. E. Houseman for this study included not only the variance of the stratified sample of acre yields, but also the variance of the estimate of wheat acreage included in the survey. Since consideration of costs in this sampling led to a change in recommendations about the appropriate design, the findings will be described in some detail. The cost function was linear in two variables: first, the expenditures associated with the sampling in a field, and second, those incident to travel between sampled fields. When cost was taken into account, it was found that because of the low cost of travel as compared with high laboratory cost of processing the samples for yield and quality, the efficiency of the sample could have been increased by 122 per cent over that of the sample as actually taken. Before the survey, little information on cost and variability was available, so the choice of interval between fields sampled and the distance to drive was made only by approximation. Upon considering cost it was found that the 122 per cent increase in efficiency accrued from taking only one unit per field rather than two, from increasing the interval between the fields sampled, and from expending the money by more than doubling the length of the miles traveled and increasing the number of fields sampled from 1,150 to 1,250. It should be explained that the advantage of increasing the length of route lies in the assumptions that total acreage is estimated by the ratio of wheat frontage to total frontage, and that the variance of this estimate is inversely proportional to the distance traveled—assumptions which have been found not unreasonable by W. A. Hendricks.

Another investigation, in character not unlike the cost study, relates to all those inquiries involving the values of two or more variates—usually many of them. Since each variate has a variance peculiar to its own distribution, to obtain the most efficient allocation of the sampling to the several strata it is usually necessary to allocate separately for every one of them. Miss M. L. Bucher is examining a compromise allocation which promises satisfactory results. This is reached by maximizing the sum of the amounts of relative information in the several variates. By relative information for each variate is meant the ratio of the information in the sample with the compromise allocation to that in a sample with optimum allocation. Neyman[4] found that this opti-

[4] "On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection," by Jerzy Neyman, *Proceedings of the Royal Statistical Society* 97: 558, 1934.

mum allocation for a single variate is that in which the number of sampling units for the ith stratum is proportional to $N_i\sigma_i$. We have, then,

$$\text{relative information} = \frac{\dfrac{1}{V_a}}{\dfrac{1}{V_o}} = \frac{V_o}{V_a} = \frac{\{S(N_i\sigma_i)\}^2/Sn_i}{S\dfrac{N_i^2\sigma_i^2}{n_i}}$$

where V_a and V_o are the variances in the compromise sample and the optimum sample, respectively, and n_i is the size of sample in the ith stratum.

The maximization of the sum of the amounts of the relative information leads to equations difficult to solve. In most cases tried, however, the third successive approximation gave satisfactorily stable results. We present in Table I the relative information in each of a group of items from two samples, listing first the relative information in the

TABLE I

PER CENT RELATIVE INFORMATION IN ACTUAL AND THEORETICAL SAMPLES

	Iowa		Indiana	
	Actual sample	Theoretical sample	Actual sample	Theoretical sample
All corn	98.36	98.80	92.56	96.06
Corn grain	97.89	98.20	92.00	96.07
Silage	81.73	81.28	70.20	82.30
Fodder	73.37	69.33	61.64	72.41
All oats	97.06	98.28	87.88	93.05
Oats grain	90.44	97.38	81.01	92.13
Barley grain	61.28	71.77	67.88	72.00
All wheat	71.40	75.65	92.10	95.08
Winter wheat grain	66.67	69.56	92.17	95.19
Summer wheat grain	58.20	60.03		
Rye grain	67.56	70.42	63.70	93.22
Popcorn	35.02	41.01		
All soybeans	94.20	90.00	85.63	94.00
Soybeans grain	89.41	66.03	77.24	87.23
Soybeans hay	98.08	95.39	91.01	97.70
Potatoes	32.51	38.07	41.99	53.60
Flax	94.21	73.78		
Alfalfa hay	97.16	97.00	91.05	86.80
Clover timothy hay	87.52	70.53	98.71	91.37
Tobacco			40.19	92.38
Cowpeas			61.29	61.01

samples as taken, and second that in samples with maximum sum of relative information. The compromise allocation produces an increase in the relative information of all of the poorly estimated items, with

corresponding decrease in only a few of those well estimated, a compromise which seems reasonable.

The value of this research lies not so much in the compromise reached, in which many items still are poorly sampled, but in the *isolation* of those items which cannot be well estimated along with the majority. The segregation of these items gives the sampler the choices of omitting them from the schedule, of proceeding with full knowledge of the small amount of information available, or of setting up separate samplings designed to elicit required information about them.

We turn now to another type of inquiry. Mr. J. R. Goodman has been working on an application of the double sampling technique in the estimation of corn yields from route survey data. The yield estimates resulting from this objective sampling are based upon two factors: (1) average number of ears per unit area, and (2) average dry weight of kernels per ear. The sampling procedure follows the principles laid down by the pre-harvest wheat surveys, with stops made at regular intervals along the route traveled. The point of entry of the field is selected as nearly as possible at random and the sampler proceeds twenty rows into the field in order to eliminate border effect. Within the field the sampling unit for the number of ears is a ten-hill area, that for size of ears (dry weight of kernels per ear) is a single hill. The ears from this one hill per field are measured, harvested, and brought to the laboratory for determination of dry weight of kernels per ear. In contrast with this expensive procedure, counts of ears can be made *at little cost; hence, three sampling units for this variate were located* in each field sampled.

The high cost of making dry weight determinations suggested the possibility of a double sampling in which ears would be measured in many fields but harvested in only a portion of them. From ear measurements taken in this and a number of earlier surveys, comparison could be made between the sampling as it was done and the proposed double sampling. It was known that the product of length times diameter of ear is highly correlated with the dry weight of kernels; hence, the dry weight of kernels per ear could be estimated by the formula

$$E_Y = \bar{y}_s + b(\bar{x}_L - \bar{x}_s)$$

where $\bar{y}_s$ is the average dry weight of kernels for the small harvested sample Y,
$\bar{x}_L$ is the average size (length$\times$diameter) of ear for the larger sample in which measurements only are taken
$\bar{x}_s$ is the average size of ear for the small sample
and b is the regression coefficient for the regression of y on x.

The variance of the estimate is given by the formula

$$V(E_Y) = \sigma_Y^2\left[(1-\rho^2)\left\{\frac{1}{n}\right\} + \frac{\rho^2}{N}\right]$$

where ρ is the correlation between the size of ear and the dry weight of kernels in the population

n = the size of the small sample

and N = the size of the large sample.

The formula[8]

$$V(E_Y) = \frac{S(Y_i - \hat{y}_i)^2}{n-2}\left[(1-r^2)\left\{\frac{1}{n} + \frac{N-n}{N}\cdot\frac{1}{n(n-3)}\right\} + \frac{r^2}{N}\right]$$

gives an estimate of the average variance in double sampling, based on the small sample values for mean deviation from regression, and using the expected value for $(\bar{x}_L - \bar{x}_s)^2$.

Since data were available for two states, Iowa and Illinois, we estimated in each of them the variance of E_Y for a double sampling in which the number of dry weight determinations, n, was 50, while the fields with ear measurements, N, was 200. In each state the relative standard error (coefficient of variation) turned out to be about $2\frac{1}{2}$ per cent. Results of the investigations are indicated in Table II. While more fields would be visited in the double sampling, only 50 dryweight determinations would be required instead of 100. The saving in laboratory costs more than compensates for the additional field measurement. Hence, it appears that double sampling may prove advantageous in the route survey work on corn yield.

TABLE II

VALUES OF n AND N FOR SAMPLINGS OF EQUAL EFFICIENCY BY TWO METHODS

Method	n Determinations of dry weight	N Ear measurements	Total fields
Simple sampling	100	0	100
Double sampling	50	200	200

The routine procedures of such agencies as the Census, the Bureau of Agricultural Economics, and the Division of Agricultural Statistics involve sampling repeated at regular or irregular intervals. Thus, the question of the efficiency of matching all or part of the sampling units in a later sampling with those in an earlier has received much atten-

[8] Developed by Professor W. G. Cochran.

tion. We have made two attacks on the problem with results that seem informative.

Jessen[9] designed his 1939 sample so that approximately half of his sampling units were drawn at random (within the strata) from those of 1938. He first considered the efficiency with which means in 1939 were estimated from this incompletely matched sample as compared with one wholly independent. The item means of the matched portion of the 1939 sample were adjusted by use of their regression on the 1938 means, and the variance of these adjusted means computed. The 1939 sample thus furnished two independent estimates of the population mean, one from the matched half and a second from the unmatched, the variances of the two estimates being different. The population mean was then estimated by weighting the means of the two halves each with the reciprocal of its own variance. For 14 items in the questionnaire, the relative efficiency of the half-matched sample as compared to one completely unmatched ranged from 122 to 145 per cent—if correlations had been perfect in this 50 per cent matched sample, the gain would have been 50 per cent for all items.

Another problem considered was that of determining the optimum allocation of sampling units among the three categories of a sampling design involving incomplete matching: (1) the first year's sample size, (2) the matched sample of the subsequent year, and (3) the unmatched sample of the subsequent year. It was assumed that the population variance was the same at the two samplings and that the sampling units were obtainable at equal and constant unit costs. For several values of the population correlation, ρ, and for a sample of 1,000 the first year, the following were the values of the matched and unmatched portions of the second year's sample:

	$\rho = \pm 0.5$	$\rho = \pm 0.9$	$\rho = \pm 0.98$
Matched	408	443	340
Independent	431	193	70
Total	839	636	410

It was found that if there is no correlation, it is a matter of indifference whether matching is done or not, but the second year's total sample must be the same size as the first, 1,000 each year. The table shows that the gain in size of total sample would be pronounced only for rather high correlations, and that some independent samples are needed no matter how high the correlation. Clearly, under the conditions of this sampling, matching close to half of the second year's units leads to a saving in cost.

[9] *Loc. cit.*

The conditions leading to this result should be emphasized. About one-third of the sampling units bore no farmsteads, so that the values of both variates in the regression were zero. This led to higher correlations than might ordinarily be encountered.

In quite a different situation, the question regarding the efficiency of matching crop acreages has been investigated by Mr. Houseman. He worked with returns from the inquiries distributed by rural mail carriers for the Agricultural Marketing Service in the Department of Agriculture. It has been the practice in the Agricultural Marketing Service to list separately the farmers reporting for two years in succession. The matching of these farms and the sending of follow-ups to those who do not report run into a considerable cost. The problem was to determine the extent to which this separate listing improves the estimate and to learn if this is an efficient way to spend the money. Houseman finds that for estimating acreages in Iowa the present method of matching farms is of practically no value. Also, because of correlation between the matched farm estimate and the more accurate ratio to land estimate which is made from the data, the matching cannot at best supply much additional information. In this kind of investigation, it is rather important that proper weights be used in combining the two estimates; otherwise, the use of the matched farms may result in a loss of information.

We have thus reached opposite conclusions about the value of matching in two instances of agricultural sampling. Both results have been somewhat surprising. Apparently the complex conditions surrounding such samplings will require many individual investigations into their efficiency.

THE MEANING OF PRODUCTIVITY INDEXES*

By W. Duane Evans and Irving H. Siegel
U. S. Bureau of Labor Statistics

Following the 1929 collapse, much attention was focussed on the subject of increasing labor productivity, mainly because of its close association with the problem of unemployment. More recently, the center of interest has shifted. Today we are striving to combine our human and material resources in such a manner as to maximize our production of war items without at the same time curtailing dangerously the manufacture of civilian essentials. Every improvement in productivity, wherever it occurs, contributes to this end. Later, when the immense task of post-war readjustment must be undertaken, progressive productivity increases may perhaps again be regarded as not an unmixed blessing. Because of these increases, we must achieve a higher scale of living than any we have ever before enjoyed, or face the prospect of persistent unemployment.

In recognition of this continuing interest in labor productivity, its background, and its implications, the Congress recently authorized the establishment in the Bureau of Labor Statistics of the Productivity and Technological Developments Division. This Division has, among its other activities, undertaken the continuation and extension of certain studies made by the National Research Project of the Work Projects Administration. It is engaged in the construction of index numbers of productivity and unit labor cost for an increasing number of industries. In addition to maintaining these quantitative historical records, it is endeavoring through field studies to ascertain unit labor requirements in selected industries. It is also collecting information on current technological and related developments which may provide some clue as to the future course of productivity.

To this new Division, employees of various branches of the government service and representatives of labor, management, and the public have been turning in increasing numbers for information and statistics relating to productivity. Through such daily contacts, we have been able to learn at first hand something about the demand side of the market for productivity indexes. We have also become aware of certain misconceptions which are prevalent among productivity-index consumers regarding the nature, proper use, and limitations of such measures. In this paper, we discuss certain points which should be taken into account in the interpretation of the types of productivity measures usually constructed.

* A paper presented at the 103rd Annual Meeting of the American Statistical Association in joint session with the American Association for Labor Legislation, New York, December 29, 1941.

A theoretical point which appears most puzzling to lay minds is the multiplicity of possible productivity indexes for, say, a manufacturing industry. Most consumers of productivity indexes approach us with the firm conviction that there must exist a unique and unequivocal productivity index referring to the particular segment of economic activity in which they are interested. To shake this conviction however so slightly, we are usually obliged to fathom the depths of our resourcefulness and, sometimes of our inquirer's patience.

Without going into the history of a controversy which once raged among statisticians themselves, we note the cardinal fact that, even with ideal data, a number of plausible productivity indexes may be constructed. Our choice is reduced to but one of these many alternatives only when the question asked about productivity change is so specific as to be equivalent to a verbalization of only one of the many possible indexes, thus precluding all indexes other than that particular one. Even if we restrict ourselves to the ordinary weighted aggregative type of index numbers, we have a whole family of permissible measures, each answering a significant variant of the same fundamental question. Each variant depends on the particular complex of quantities chosen for the products included in the productivity index, and the number of possible quantity complexes, real and hypothetical, is without limit.

In reply to a more sophisticated inquirer, we may resort to a symbolic demonstration. To simplify the discussion, we turn from productivity to the unit labor requirement, which is the reciprocal of productivity. For a single uniform product, the unit labor requirement for any particular time is the labor consumed per unit of output of the specified good; or, what amounts to the same thing, the ratio of the total labor required for the production of a given volume of a homogeneous good to the given volume of that good.[1]

The unit-labor-requirement index for a single homogeneous product is merely the ratio of the unit labor requirements for the times t_i and t_0, where t_0 denotes the base period.[2] Symbolically, this may be expressed as follows, where I_r represents the index of unit labor require-

[1] The preferred measure of labor is generally man-hours rather than the number of men. As is pointed out later, all labor time is of necessity regarded as additive; the nature of the available statistics usually does not permit any distinction to be made between hours of different intensity or hours of different kinds of workers.

[2] It should be noted that productivity indexes as usually constructed reflect fluctuations in the degree of capacity utilization, as well as influences of a more permanent character. For various purposes, it would be desirable to construct productivity indexes which would be independent of changes in the degree of capacity utilization. Such indexes by inference would involve, in both numerator and denominator, estimates of unit labor requirements at the same operating level. The unit labor requirements would be wholly hypothetical if the operating level assumed for the comparison differed from that actually attained. Since, in general, the data available do not permit the construction of such measures, the discussion in the text is limited to the more usual indexes.

ments, I_p represents the productivity index, and r_i and r_0 denote the unit labor requirements at times t_i and t_0, respectively:

$$I_r = \frac{1}{I_p} = \frac{r_i}{r_0} = \frac{q_x r_i}{q_x r_0} = \frac{m_{xi}}{m_{x0}}. \tag{1}$$

It may be observed that the index is also equivalent to the ratio of the total amount of labor required to produce any specified volume of the good in question at t_i to the total required for the same volume at t_0. This has been shown in equation (1) by introducing q_x, which represents the specified volume, and m_x, which denotes the required amount of labor. It should be noted at this point that, whatever the quantity q_x, the index for one product is invariant.

For the more important case of an industry producing a number of products, we need merely relate the total volume of labor required to produce a given complex of quantities at $t_i(M_{xi})$ to the total volume required for the same complex at $t_0(M_{x0})$. We may represent such an index for n goods as follows:

$$I_r = \frac{M_{xi}}{M_{x0}} = \frac{\sum^n m_{xi}}{\sum^n m_{x0}} = \frac{\sum^n q_x r_i}{\sum^n q_x r_0}. \tag{2}$$

The totals are, of course, equivalent to the sums of the totals for the individual products.

Now, there is a very important difference between the labor requirement (or productivity) index for a group of products and the labor requirement (or productivity) index for a single product. An index for one product, as we have already observed, is invariant since it depends in no way on the quantity of output assumed. An index for many products, however, does vary according to the quantity complex assumed. Now, the quantities for t_i (i.e., the $q_x = q_i$) generally give a different result from that given by the quantities for t_0 (i.e., the $q_x = q_0$); again, hypothetical quantities may give still different results. In some cases, the differences are unimportant; in others, they may be considerable. Which of the infinite number of possible indexes is correct? It all depends on the quantity complex one has in mind. If one does not specify a particular quantity complex, none of the indexes is necessarily best, and there is no basis for choosing one above all the others.

The bald question, "How much has productivity increased in a given multi-product industry?" may be compared with the question, "How much larger is one house than another?" A variety of answers to the latter question might be given—depending on whether the number of

rooms, the cubic space enclosed, the ground area covered, or the construction cost is adopted as a criterion. In the same way, the former question has no single numerical answer unless a particular quantity complex is specified as the desired criterion.

We are frequently asked to appraise productivity indexes constructed from actual data. Almost invariably, the formulas used differ from those we have discussed above and the methods are less direct. The available data, moreover, are not complete, nor are they perfect within their restricted compass. The net result of compromise, assumption, substitution, ingenuity—and, sometimes, vagueness of purpose—is seldom ascertainable.

It is essential before attempting to evaluate an actual productivity index to know what formula is being approximated. The selection of a formula, however, is not often a "conscious" act. Availability of data is usually the factor which determines the choice.

Equation (3a) represents the general form of a proper productivity index. In it, the quantities of the individual goods involved, the q_s, may be real, referring to actual production in any period of time, or hypothetical. In equations (3b) and (3c), the quantity complexes are those of the base period, t_0, and the index period, t_i, respectively. To many, these latter forms appear inherently more "reasonable" than those involving other quantity complexes.

$$I_{ps} = \frac{1}{I_{rs}} = \frac{\sum^{n} q_s r_0}{\sum^{n} q_s r_i} \tag{3a}$$

$$I_{p0} = \frac{\sum^{n} q_0 r_0}{\sum^{n} q_0 r_i} \tag{3b}$$

$$I_{pi} = \frac{\sum^{n} q_i r_0}{\sum^{n} q_i r_i} \tag{3c}$$

The form indicated by equation (3c) is the one usually chosen, consciously or otherwise, for actual evaluation of a productivity index. The evaluation of equation (3a) or (3b) implies knowledge of the unit labor requirements (or numbers proportionate to them) for all the goods involved for each of the years. In contrast, the individual unit labor requirements for only the base year are needed for equation (3c); the denominator may generally be obtained directly as total man-hours of labor expended in the period t_i.

Estimates of the unit labor requirements in the base period are sometimes available, perhaps from the results of sample field studies. Usually, however, the available figures do not cover all the goods produced by an industry. If the r_0 cannot satisfactorily be estimated for all products, equation (3c) will not give an index of unity for the base period. It may, however, be adjusted as follows to achieve this result:

$$I_{pt}' = K \frac{\sum^{n-z} q_t r_0}{\sum^{n} q_t r_t} = \frac{\sum^{n} q_0 r_0}{\sum^{n-z} q_0 r_0} \cdot \frac{\sum^{n-z} q_t r_0}{\sum^{n} q_t r_t}. \tag{4}$$

Here the symbol z is used to denote the products for which unit labor requirements in the base period are not available.

Now, equation (4) may be rewritten as follows:

$$I_{pt}' = \frac{\sum^{n-z} q_t r_0}{\sum^{n-z} q_0 r_0} \Big/ \frac{\sum^{n} q_t r_t}{\sum^{n} q_0 r_0}. \tag{5}$$

In this form, the numerator is a production index (aggregative form) for the several products with r_0 weights. The denominator is the index of total man-hours for all the products. The quotient form of a productivity index is very common, and most actual productivity indexes are derived in this form.[3]

But our retreat from precision toward approximation does not end here. Frequently, estimates of the r_0 are not available for even a few products. Consequently, those who construct production indexes to be used in measuring productivity generally substitute other weighting factors more or less related to unit labor requirements. Unit labor costs are regarded as a first approximation, since they will be proportional to unit labor requirements if average hourly earnings in the production of the several goods considered are the same.[4] These weights also are seldom available, so value added by manufacture per unit of product, if it is available, may be used as a second approximation. In this case, the

[3] Usually, the production and man-hours indexes are not comparable in scope for the additional reason that the quantities entering into the production measure refer to output of *all* establishments while the man-hours index is restricted to the establishments in an "industry." Some investigators, like Dr. Fabricant (in *Output of Manufacturing Industries: 1899–1937*) prefer to make a formal value adjustment by the use of Dr. Mills' "value-adequacy ratio" or its equivalent. By such an adjustment, they hope to correct for the omission of the z products from the production index and for all other differences in scope between the production and man-hours indexes.

[4] In this paragraph and the following, a discussion of the condition of equivalence of indexes with different sets of weights is, for simplicity, restricted to the case of proportionality. Proportionality, however, is merely a sufficient and not a necessary condition; it may be shown, for example, that equality results whenever certain correlation coefficients reduce to zero. (See I. H. Siegel, this JOURNAL, September 1941, pp. 343–350, and December 1941, pp. 519–521.)

additional assumption is made that wages constitute the same proportion of value added by manufacture for each of the products. Unit value is the next step in the line of approximations; the validity of its use depends, in addition to all prior assumptions, on the constancy from product to product of the ratio between value and value added. Unit value is probably the most commonly used weighting factor, since it is frequently available for the more important products of an industry from the Census of Manufactures. When unit values are lacking, recourse may be had to wholesale prices as weights. These, of course, will be influenced by the additional factor of distribution costs.

It thus comes about that production indexes constructed for measuring productivity changes are frequently like indexes constructed for measuring production *per se;* the latter often are of the aggregative form and also embody pecuniary weights, though by choice rather than necessity. Hence, a much more usual approximation than (5) is the following:

$$I_{p1}'' = \frac{\sum^{n} q_1 v_s}{\sum^{n} q_0 v_s} \Big/ \frac{\sum^{m} q_1 r_1}{\sum^{m} q_0 r_0} \qquad (6)$$

where the v_s are substitute weights, usually unit values for one or more years.

The widespread and unfortunate assumption that a productivity index can be satisfactorily obtained by relating any production index to a man-hour index probably has its origin in the form of equation (5). Needless to say, unless we are to revise our basic notions of labor productivity, the result will be valid only when the production index employed is a satisfactory approximation to the proper aggregative production index with labor-requirement weights for all products.

The approximative formulas given above are by no means exhaustive. Some of the production indexes actually used in constructing productivity measures are not of the weighted aggregative type, but, rather, are geometric means of relatives or are of the unweighted aggregative type. Some indexes are based on substitutes for output altogether, such as sales, shipments, or consumption. Finally, there are many problems of chronological continuity, which necessitate various adjustments; such adjustments, in effect, make the formulas actually used more complicated than those here shown.

The validity of an approximative index of the kind discussed here depends upon how closely it conforms to the "true" index, which, explicitly or implicitly, is as defined in equation (3c). The degree of similarity, of course, is usually unknown. If it could be ascertained in some

cases, it might provide some clue to the error to be expected in other cases where no direct comparison of the true and approximative indexes is possible. Thus, it is of interest to know whether the production index in (5) for a few products with r_0 weights adequately represents the index for all or most products with r_0 weights. It is also of interest to know whether substitute weights of the kind used in (6) really tend to be proportional to the r_0, so as to give us a close approximation to (3c).

The WPA National Research Project undertook, in the course of constructing its indexes, to test approximative methods wherever possible. The opportunities were few and the tests by no means definitive. The results indicated that approximative indexes were adequate over long periods but not so trustworthy for year-to-year changes. In extreme instances, the unit labor requirements may tend to be inversely related to the substitute weights. In such cases, the derived productivity index may even lie outside the range of the productivity relatives for the individual products. No "directly" computed productivity index would give such a result, for such an index is always equivalent to an average of the productivity relatives for the individual products.

A large demand exists for monthly or other short-term productivity indexes. Unfortunately, many obstacles lie in the way of either interpreting such short-term measures properly or even obtaining satisfactory data in the first place. Consequently, we are reluctant to construct short-term productivity indexes and hesitate to "sanction" any referred to us for criticism.

The difficulties which beset the construction of correct annual indexes also beset the construction of indexes for short periods. But there are other difficulties peculiar to short-term indexes which are even more serious. For example, the cycle of production may be longer than the —say, monthly—period for which the index is constructed; consequently, the labor expended in any month may correspond only partly to the output reported for that month. Furthermore, since it is difficult to obtain production series on a short-term basis, resort is very frequently had to substitute quantity series such as shipments or consumption.[5] These substitute series may give an absolutely false indication of the short-term movements of the true weighted output. They may lag behind or lead true production; the magnitude and direction of the changes they indicate may be spurious, especially if there is storage between stages of production and the cycle of fabrication is longer than one month.

Even if the available data are in satisfactory form, the significance

[5] In some instances, for want of something better, man-hours series themselves are used as indicators of production.

of short-term productivity movements may be obscure. If an industry produces different goods through the year, the month-to-month movements may reflect differences in the composition of the output, as well as true changes in level. In any industry, and particularly in seasonal ones, monthly productivity changes may be dominated by changes in the degree of capacity utilization. In such cases, it would indeed be difficult to disentangle seasonal variations from longer-term trends—except in retrospect. Thus, it may be said categorically that any monthly productivity index constructed from currently available data should be used with caution. Preliminary investigations, however, indicate that in some cases indexes constructed on a less-than-annual basis may convey useful information not revealed by simple annual averages.

The nature of an index of productivity is often misunderstood. Those who regard it as a touchstone are frequently chagrined to find that it reports the base alloy together with the gold. Some of our inquirers seem to believe that, since such an index is frequently called a "labor productivity" index, it reflects changes in output per unit of labor which are attributable principally to essential improvements in labor efficiency. Others tend to interpret the changes as resulting from changes in technology alone. Neither of these views is correct. Increases in output per unit of labor reflect the joint efforts of all factors of production; sometimes one factor may dominate and sometimes another.

As has already been indicated, non-technological factors, such as changes in level of capacity utilization and changes in the composition of production, may influence the course of productivity. It may, perhaps, be considered a defect of available statistics that they do not permit us to construct measures which would reveal only changes which are of a permanent character. Capricious fluctuations in productivity resulting from changes in capacity utilization are of an impermanent character and reflect no change in the state of production techniques.

It is important also not to overlook the fact that, whatever the reason for a change in the productivity level of an industry, the index is an average and subject to the limitations of all averages. Thus, an industry index should not be expected to depict the change occurring in any particular plant or branch of an industry. The mere shifting of production from plants at one level of productivity to plants at another may affect the average level for an industry as a whole.

It is frequently assumed that rising productivity necessarily implies labor displacement. Though it sometimes does involve direct labor displacement, it may also have less drastic and more subtle implications. Indeed, it may even go hand in hand with rising employment, as

it has in the rayon industry during the 1930's and as it has during the expansive phases of most, if not all, other manufacturing industries.

There is a fundamental difference between production *per se* and production as implied in a productivity index that is frequently overlooked. In many production indexes, final output is the only matter of interest. In connection with productivity, on the other hand, production is understood to imply the movement of specified materials through a given series of manufacturing operations. For example, a trend toward the purchase of semi-finished materials by an industry might properly be ignored in an index of final output. In a productivity index, on the other hand, a spurious increase, unrelated to any actual technological change, might result from such a trend, unless appropriate allowance were made for it.

It may also be observed in passing that there is a fundamental assumption which is seldom explicitly recognized involved in the compilation of all the common productivity indexes. It is assumed that all human labor is interchangeable and additive, and that no more weight is to be attached to an hour of skilled labor than to an hour of unskilled labor in building up man-hour totals. A technological change which merely permits the substitution of lower- for higher-paid labor does not affect a productivity index.

From what has been said, it is clear that much unfinished business remains in the field of productivity indexes. Perhaps the major need is for unit-labor-requirements data covering a greater number of products. Such information, which is of intrinsic interest, is also essential to the construction of more satisfactory productivity indexes. The Bureau of Labor Statistics, through field studies, will continue to collect such data. It should be stressed that a substantial contribution to the available body of knowledge of labor requirements could be made by persons in industry or by investigators in the academic field. There has been in recent years a substantial improvement in both production and labor series with respect to adequacy of coverage and the degree of detail shown. A continuation of this trend will make possible the construction of more adequate productivity measures. It is evident from the discussion in this paper that improvements in the available materials will call forth the application of more refined statistical techniques.

MACHINE METHOD FOR THE EXTRACTION OF CUBE ROOT

By Otto E. Lancaster
University of Maryland

In a recent study of sequences defined by algebraic difference equations[1] the author found that a sequence of numbers defined by the recurrent relation

$$A_{n+1} = \left(\frac{A_n^3 + 2B}{A_n^3 + .5B}\right)\frac{A_n}{2} \tag{1}$$

converges very rapidly to the cube root of B. That is, if A_1 is an approximation to the cube root of B, then

$$A_2 = \left(\frac{A_1^3 + 2B}{A_1^3 + .5B}\right)\frac{A_1}{2}$$

is a better approximation;

$$A_3 = \left(\frac{A_2^3 + 2B}{A_2^3 + .5B}\right)\frac{A_2}{2}$$

is a still better approximation; etc. These numbers approximate to the cube root of B much more rapidly than the corresponding numbers defined by the algorithm,

$$A_{n+1} = \frac{B/A_n^2 + 2A_n}{3}, \tag{2}$$

which is the basis of the present machine method of calculation. If two consecutive numbers A_n and A_{n+1} of a sequence defined by (1) agree in the first i digits, then A_{n+1} is correct to $3i-2$ digits. Whereas if two consecutive numbers of a sequence defined by (2) agree in the first i digits, then the second number is correct to $2i-1$ digits.

It is the purpose of this paper to outline a procedure to use on a calculating machine.[2] The procedure given here is adapted to a ten place Monroe calculating machine equipped with automatic division and semi-automatic multiplication (use of plus and minus bars). Some variations in technique may be necessary if other calculating machines, or even other Monroe models, are to be used.[3]

[1] "Sequences Defined by non-linear Algebraic Difference Equations," *Annals of Mathematics*, Vol. 42, 1941, pp. 251–280.

[2] This is published with special permission of the Monroe Calculating Machine Company.

[3] The author desires to express his thanks to Mr. Fred C. Diedrich for his suggestions.

First, commence at the decimal point and mark the number off into periods of three digits each. Shift the decimal point of the number so that it lies between the digits of the first and second period. (The decimal point of the answer can then be placed in the proper place after the calculation is completed.) Use the following decimal set-up

upper-dials	7
keyboard	7
lower-dials	14.

First Approximation. Mentally approximate the root. Call this value A_1. Particular care need not be used in obtaining this estimate. However, the closer the approximation, the fewer operations needed.

Second Approximation. 1. Set A_1 on the keyboard and multiply by A_1. Copy the result on the keyboard (depress minus bar to prove correct transfer). Clear dials and multiply by A_1 to produce the cube of the approximation.

2. Set the base number, B, on the keyboard. Multiply by .5000000. Mentally note the sum of the products (numbers on lower dials). Clear upper dials only.

3. With base number still on the keyboard multiply by 1.5000000 to produce numerator in formula. Clear upper dials and keyboard.

4. Set sum mentally noted in step 2 on keyboard.[1] Line up carriage for division and push division lever. Clear keyboard and lower dials.

5. Set A_1 on the keyboard and return upper dials to zero with plus bar. Enter 2.0000000 on the keyboard and push division lever. The second approximation appears in upper dials. Clear lower dials and keyboard. Record A_2.

Third Approximation. 1. Set the first three digits of A_2 on the keyboard and return the first three digits (on left) of upper dials to zero with plus bar. Copy result in lower dials on keyboard and depress minus bar to prove correct transfer. Multiply amount on keyboard by the first three digits of A_2. Clear upper dials and keyboard.

Continue as in other steps of Second Approximation.

If A_3 and A_2 agree in the first three digits, then the first seven digits of A_3 are correct. (Note this does not mean that the result is correct to seven digits for the machine does not round the numbers and the 8th digit may be 5 or more. Moreover, it may be incorrect.) If A_3 and A_2 agree in the first two digits, then the first 4 digits of A_3 are correct. If one's first approximation is the nearest half integer, the third approximation is correct to the first seven digits. No more correct digits can

[1] If this sum contains four digits to left of the decimal point, shift the decimal point on the keyboard one place to the right for the two operations in steps 4 and 5.

be assured with the above decimal set up. (By a modification, the method may be used to obtain the first nine correct digits.)

EXAMPLE

Consider the example given in Monroe's pamphlet on "Machine Methods for the Extraction of Cube and Other Roots."

$$\sqrt[3]{128.53}$$

First approximation, let $A_1 = 5.0000000$.

Second Approximation. 1. Set 5.0000000 on the keyboard and multiply by 5.0000000. Copy result in the lower dials on the keyboard and depress the minus bar to prove the correct transfer. Clear the upper dials. Multiply the amount on the keyboard by 5.0000000 to produce the cube of the approximation. Clear the upper dials and the keyboard.

2. Set the base number 128.53 on the keyboard and multiply by .5000000. Mentally note the sum of the products, 189.265. Clear upper dials only.

3. With the base number still on the keyboard multiply by 1.5000000. Clear the upper dials and the keyboard.

4. Enter the number noted mentally in step 2 on the keyboard. Line up carriage for division and push the division lever. (The result is 2.0180510.) Clear the keyboard and the lower dials.

5. Set the approximation 5.0000000 on the keyboard and return upper dials to zero with plus bar. Set 2.0000000 on the keyboard and push the division lever. The second approximation, 5.0400275, appears in the upper dials. Record this result. Clear the lower dials and the keyboard.

Third Approximation. 1. Set 5.04 on the keyboard and return the first three digits (5.04) of the upper dials to zero with the plus bar. Copy the result in the lower dials on the keyboard and depress minus bar to prove the correct transfer. Clear the upper dials. Multiply the amount on the keyboard by 5.04. Clear the upper dials and the keyboard.

2. Set the base number, 128.53, on the keyboard and multiply by .5000000. Mentally note the result (192.289064). Clear the upper dials.

3. With the base number still on the keyboard, multiply by 1.5000000. Clear the upper dials and the keyboard.

4. Enter the number 192.289064 mentally noted in step 2 on the keyboard. Push the division lever. Clear lower dials and keyboard.

5. Set 5.04 on the keyboard and return the upper dials to zero with the plus bar. Clear the keyboard. Set 2.0000000 on the keyboard and push the division lever. The third approximation 5.0460303 appears in the upper dials. Since the first three digits are the same as 5.04, the

result is correct for the first seven digits. Since occasionally one makes an error in the use of the machine the result should be checked.

The advantage of this method over the old method is more readily illustrated by examples which are not so nearly equal to a perfect cube.

A similar rule for obtaining other roots, based on the general relation

$$A_{n+1} = \left(\frac{(p-1)A_n^p + (p+1)B}{(p+1)A_n^p + (p-1)B}\right)\cdot A_n$$

could be given. This formula becomes

$$A_{n+1} = \left(\frac{A_n^p + \dfrac{p+1}{p-1}B}{A_n^p + \dfrac{p-1}{p+1}B}\right)\cdot\left(\frac{p-1}{p+1}\right)\cdot A_n,$$

when it is written in an analogous way to the formula for cube root. However, the author does not deem it applicable at present.

JAMES WATERMAN GLOVER, 1868-1941

The death of James W. Glover on July 15, 1941, was the end of the career of a man who contributed as much to the development of instruction in the mathematics of finance, actuarial science, and mathematical statistics in American universities as any other individual.

He was a member of the Mathematics Department of the University of Michigan from 1895 to 1938, the date of his retirement, becoming a full professor in 1911 and later serving as chairman of his department. In 1902 he began instruction in the mathematics of finance and insurance and under his direction there soon developed the broad and sound curriculum in actuarial science for which Michigan became noted. For many years the annual examinations of the American actuarial societies have been given in Ann Arbor. He was one of the first in America to foresee the coming importance of mathematical statistics and very soon after the beginning in actuarial science, instruction in the related but broader subject was begun. Professor H. C. Carver joined the staff in 1916, and he was actively encouraged by Professor Glover to develop the work in mathematical statistics which has since come to occupy so prominent a place at the University of Michigan and which has added to its reputation. An early feature of instruction in these two subjects was the provision of calculating machines for the use of all the students; in particular, when Angell Hall was built in 1925, Professor Glover took the opportunity to include in it statistical laboratories equipped in a way unique at that time. Excellent libraries in these two fields were built up. As an administrator Professor Glover was successful, not only strengthening the staff in the subjects he was most directly concerned with, but also in a notable way in pure mathematics.

Professor Glover's own work was largely in actuarial science and population statistics. As early as 1906 he was consultant to the Canadian Royal Commission on Insurance. In 1906 and 1907 he was consulting actuary to the Wisconsin Legislative Committee and to the Joint Committee on Banks and Insurance. These were typical of a long career as an insurance and statistical expert for many public and private agencies. One must mention his long association with the U. S. Census Bureau, his special service to his University and his state with respect to the setting up of pension funds, and his extended participation in the affairs of the Teachers Insurance and Annuity Association which he served as president during the two years, 1930-1932.

In addition to many articles, his longer published works include the especially noteworthy *U. S. Life Tables for 1890, 1901, 1910, and 1901–10* (Government Printing Office), the widely used *Tables of Applied Mathematics in Finance, Insurance and Statistics* (George Wahr, Ann Arbor, 1923), and the *Mathematics of Life Insurance* (with W. O. Menge) (Macmillan, 1935).

He was long a member of the American Statistical Association and was elected a Fellow about 1917.

CECIL C. CRAIG

PROCEEDINGS

103RD ANNUAL MEETING

BILTMORE HOTEL, NEW YORK

PROGRAM

Saturday, December 27, 1941

—10:00 A.M.—

STATISTICAL EXAMINATION OF INVESTMENT EXPERIENCE
Chairman: Donald B. Woodward, Mutual Life Insurance Company of New York

Statistical Examination of Mortgage Experience
Part 1. Lawrence N. Bloomberg, United States Housing Authority
Part 2. Mortimer Kaplan, Federal Housing Administration
Discussion: Frank J. Travers, Lincoln National Life Insurance Company
Clifford C. Boyd, Federal Home Loan Bank Board

Statistical Examination of Corporate Bond Experience
Harold G. Fraine, Securities and Exchange Commission
Discussion: O. Paul Decker, American National Bank and Trust Company

Research in Municipal Finance
Harry L. Severson, Federal Deposit Insurance Corporation
Discussion: Susan S. Burr, Board of Governors of the Federal Reserve System
Sanders Shanks, Jr., *The Bond Buyer*
C. A. Sienkiewicz, Federal Reserve Bank of Philadelphia

ROLE OF A DIVISION OF RESEARCH AND STATISTICS IN AN ADMINISTRATIVE AGENCY
Chairman: F. Leslie Hayford, New York City
In a Business Organization
Henry B. Arthur, Swift & Company
In the Field of Relief and Social Work
Robert T. Lansdale, Community Service Society
In the Federal Government
Woodlief Thomas, Board of Governors of the Federal Reserve System
Discussion: Donald Belcher, American Telephone and Telegraph Company
Ewan Clague, Social Security Board

Saturday, December 27, 1941

— 10:00 A.M. —

Impact of Defense on Special Economic Areas

Chairman: Theodore O. Yntema, University of Chicago and Cowles Commission

Impact of War upon Industrial Capacity and Investment

M. Joseph Meehan, Bureau of Foreign and Domestic Commerce

Impact of Defense on Labor Supply and Labor Utilization

Louis Levine, Social Security Board

Effects of a Relatively Stationary Population Total

Round Table

Chairman: Edmund deS. Brunner, Teachers College, Columbia University

Discussion: Richard M. Bissell, Jr., Harvard University
Louis I. Dublin, Metropolitan Life Insurance Company
Frank W. Notestein, Princeton University
Leon E. Truesdell, Bureau of the Census
Clyde V. Kiser, Milbank Memorial Fund
Frank Lorimer, American University
P. K. Whelpton, Scripps Foundation for Research in Population Problems
Alan Sweezy, Williams College
O. E. Baker, Bureau of Agricultural Economics

— 2:30 P.M. —

Role of Tests of Significance in Biological Research

(Biometrics Section with the Institute of Mathematical Statistics)

Round Table

Chairman: Edwin B. Wilson, School of Public Health, Harvard University

Discussion: W. Edwards Deming, Bureau of the Census
Harold Hotelling, Columbia University
Lowell J. Reed, Johns Hopkins University School of Hygiene and Public Health
George W. Snedecor, Iowa State College

Selected Problems in Fiscal Policy

(With the American Economic Association)

Chairman: W. L. Crum, Harvard University

The Theory of Public Investment

Seymour E. Harris, Harvard University

The Capital Budget Idea in the Light of the Defense Program

Morris A. Copeland, Office of Production Management

Accounting for Public Investment

George O. May, Price, Waterhouse & Company

Discussion: Dan T. Smith, Harvard University
Rufus S. Tucker, General Motors Corporation

Saturday, December 27, 1941

—2:30 P.M.—

STATISTICS IN RELIEF ADMINISTRATION

(Arranged by the Joint Committee on Relief Statistics)

Chairman: Ralph G. Hurlin, Russell Sage Foundation

Measurement of the Effect of Defense Employment on Relief Loads
Walter Perkins, Social Security Board
Validity of Estimates of Employability of Relief Recipients
Robert P. Wray, Pennsylvania Department of Public Assistance
Measurements of the Normality of Relief Populations
Raymond Willoughby, Rhode Island Department of Social Welfare
Discussion: Emmett H. Welch, Bureau of Labor Statistics
Irma Rittenhouse, New York State Department of Labor

THE USE OF CENSUS DATA IN ANALYZING MARKETING PROBLEMS

(With the American Marketing Association)

Chairman: Vergil D. Reed, Bureau of the Census

Small Area Census Data and Their Use in Marketing Analysis
Alfred N. Watson, Curtis Publishing Company
Recent Changes in Food Marketing as Revealed by the Census
Carl W. Dipman, *The Progressive Grocer*
Recent Changes in Industrial Markets Revealed by the Census
Lowell J. Chawner, Bureau of Foreign and Domestic Commerce
Discussion: Paul Nystrom, Columbia University

ROLE OF A DIVISION OF RESEARCH AND STATISTICS IN AN ADMINISTRATIVE AGENCY

Round Table

Chairman: Meredith B. Givens, New York State Department of Labor

—8:00 P.M.—

CRITERIA IN ESTABLISHING PRICES BY THE GOVERNMENT

(With the American Farm Economic Association)

Chairman: Frederick C. Mills, Columbia University

Criteria in Establishing Farm Prices
Warren Waite, University of Minnesota
Criteria in Establishing Non-Agricultural Prices
Donald Wallace, Office of Price Administration
Discussion: Clayton Gehman, Board of Governors of the Federal Reserve System
Sidney Hoos, Commodity Credit Corporation

Saturday, December 27, 1941

— 8:00 P.M. —

CONTRIBUTED PAPERS BY YOUNGER MEMBERS

Chairman: W. L. Crum, Harvard University

Selective Service's Medical Statistics Program
Oliver H. Folk, Selective Service System

Conspicuous Consumption as Measured by Clothing Expenditures
A. J. Jaffe, Bureau of the Census

The Method of Classifying Areas with Respect to Housing Characteristics
Beatrice H. Lapin, Office of Defense Housing Coordinator

Gross and Net Unemployment: The Importance of Definition in the Measurement of Unemployment
Clarence D. Long, Jr., The Institute for Advanced Study

The Limitations of Tests of Significance for Index Numbers of Prices
Nilan Norris, Hunter College

A Critical Appraisal of Business Statistics
John S. Perkins, Boston University

Sunday, December 28, 1941

— 9:00 A.M. —

BREAKFAST MEETING—OFFICERS, DIRECTORS, AND DISTRICT REPRESENTATIVES

— 10:00 A.M. —

NEW METHODS OF NUMERICAL CALCULATION

I. Statistical and Matrix Calculation
(With the Institute of Mathematical Statistics, and with the cooperation of the Committee on Addresses in Applied Mathematics of the American Mathematical Society)
Technical Session

Chairman: Charles R. Langmuir, Carnegie Foundation for the Advancement of Teaching

Some New Matrix Methods in Least Square and Other Multivariate Problems
Harold Hotelling, Columbia University

The Mallock Electrical Calculating Machine for Solving Simultaneous Linear Equations
Elizabeth Monroe Boggs, Cornell University

Mathematical Operations with Punched Cards
J. C. McPherson, International Business Machines Corporation

Recent Developments in Correlation Technique
Paul S. Dwyer, University of Michigan

Sunday, December 28, 1941

—2:30 P.M.—

CONTRIBUTED PAPERS I

(Arranged by the Biometrics Section)

Chairman: Hugo Muench, Rockefeller Foundation

Energy at the Threshold of Vision
Selig Hecht, Columbia University

Some Aspects of the Multiple Hit Problem
C. P. Winsor and J. W. Tukey, Princeton University

Biologic Analysis in Allergy
Margaret R. Pabst, Work Projects Administration

Environment Effects on Sex Ratio, Malformations and Habitus of the Newborn
W. F. Peterson, University of Illinois Medical School, and Alvin Mayne, Office of Production Management

PERSONNEL SELECTION AND VOCATIONAL GUIDANCE

Chairman: Harry A. Jager, U. S. Office of Education

Validating Employment Tests
C. L. Shartle, Social Security Board

Personnel Selection in the Public Service
Albert H. Aronson, Social Security Board

A Study of Vocational Adjustment of Metropolitan Youth
Irving Lorge, Teachers College, Columbia University

Discussion: Albert K. Kurtz, Life Insurance Sales Research Bureau

NEW METHODS OF NUMERICAL CALCULATION

II. Mechanical Solution of Differential Equations

(With the Institute of Mathematical Statistics, and with the cooperation of the Committee on Addresses in Applied Mathematics of the American Mathematical Society)

Technical Session

Chairman: Ronald M. Foster, Bell Telephone Laboratories

Punch-card Calculation of Orbits
W. J. Eckert, Naval Observatory

Punch-card Methods for Solving Linear Differential Equations of Second Order
Martin Schwarzschild, Columbia University

Differential Analyzers
S. H. Caldwell, Massachusetts Institute of Technology

Discussion: L. S. Dederick, Aberdeen Proving Ground
Norbert Wiener, Massachusetts Institute of Technology

Sunday, December 28, 1941

2:30 P.M.

CONSUMER PREFERENCE STUDIES

(With the American Marketing Association)

Chairman: Lawrence C. Lockley, Curtis Publishing Company

Psychological Difficulties in Measuring Consumer Preferences
Albert Blankenship, N. W. Ayer & Son
Sales Tests as a Method of Product Testing
Ross Cunningham, Massachusetts Institute of Technology
Pre-Testing of Products by Consumer Juries
Mason Smith, General Foods Corporation
Reaction of Farmers to Agricultural Programs
Rensis Likert, Department of Agriculture
Techniques of Appraising Brand Preference by Consumer Interviewing
Harry D. Wolfe, Kent State University
Discussion: Ralph Cassady, Jr., University of California
Oliver F. Benz, duPont Company

WAGES, PRICES, RENTS, AND INCOMES

Chairman: Aryness Joy, Bureau of Labor Statistics

Measurement of Wage Differentials in Manufacturing Industries
Robert Myers, Bureau of Labor Statistics
The Problems of Determining Fair Rents
Karl Borders, Office of Price Administration
Prices and Wages—Recent Trends
Walter J. Keim, Bureau of Labor Statistics

—6:30 P.M.—

DINNER MEETING OF THE BIOMETRICS SECTION, FOLLOWED BY THE BUSINESS MEETING OF THE SECTION

—8:00 P.M.—

THE CHANGING POSITION OF THE BANKING SYSTEM AND ITS IMPLICATIONS FOR MONETARY POLICY

(With the American Economic Association)

Chairman: Oliver M. W. Sprague, Harvard University

The Implications of Fiscal Policy for Banking and for Monetary Control
John H. Williams, Harvard University
The Future of Commercial Banking
Murray Shields, Irving Trust Company
The Future of the Government Loan Agencies
Neil H. Jacoby, University of Chicago

Sunday, December 28, 1941

—8:00 P.M.—

On Some Technical Aspects of Sampling

(With the Institute of Mathematical Statistics)

Technical Session

Chairman: Helen M. Walker, Teachers College, Columbia University

On the Relative Efficiencies of Various Areal Sampling Units in Population Inquiries

Morris H. Hansen and William Hurwitz, Bureau of the Census

On the Monthly Sample Survey of Unemployment

Lester R. Frankel and J. Stevens Stock, Work Projects Administration

On Certain Biases in Samples of Human Populations

J. Cornfield, Bureau of Labor Statistics

On the Relation of Probability to Sampling

William G. Madow, Bureau of the Census

Recent Developments in Sampling for Agricultural Statistics

George W. Snedecor and Arnold J. King, Iowa State College

Discussion: William G. Cochran, Iowa State College
Joseph A. Greenwood, Duke University
Meyer A. Girshick, Bureau of Agricultural Economics

Monday, December 29, 1941

—9:00 A.M.—

Annual Business Meeting. Reports and Election of Officers

—10:00 A.M.—

Measuring National Defense

Chairman: Henry B. Arthur, Swift & Company

Measuring National Income as Affected by the War

Milton Gilbert, Bureau of Foreign and Domestic Commerce

Measuring Economic Burdens of Defense in Various Nations

Raymond Goldsmith, Securities and Exchange Commission

Measuring Defense Progress

Lester Kellogg, Office of Production Management

Savings and Investments and the Emergency

(With the American Finance Association)

Chairman: C. O. Hardy, Brookings Institution

National Income, Savings, and Investments Since 1880

Simon Kuznets, National Bureau of Economic Research

Savings and Investments as Reflected in Corporation Reports

Rexford C. Parmelee and Irwin Friend, Securities and Exchange Commission

Discussion: Claude L. Benner, Continental American Life Insurance Company
George Terborgh, Machinery and Allied Products Institute
Richard V. Gilbert, Office of Price Administration

Monday, December 29, 1941

—2:30 P.M.—

PROBLEMS OF PRICE CONTROL

(With the American Economic Association)

Round Table

Chairman: Don D. Humphrey, Office of Price Administration

EXPERIMENTS ON INSECT CONTROL, THEIR DESIGN AND ANALYSIS

(Arranged by the Biometrics Section with the cooperation of the American Association of Economic Entomologists, Eastern Branch)

Chairman: F. Z. Hartzell, New York Agricultural Experiment Station

Comparison of the Toxicity to Aphis rumicis *of Certain Organic Compounds*
William Moore, American Cyanamid Company
Effect of Variable Copper-lime Bordeaux Ratios in Flea Beetle Control
W. A. Rawlins, Cornell University
Relation of Corn-borer Biology to Field Plot Technique
Neely Turner, Connecticut Agricultural Experiment Station
Plot Design and the Combination of Observations in Insecticidal Experiments on the Codling Moth
H. N. Worthley, Pennsylvania State College

—8:00 P.M.—

CLINIC ON ENTOMOLOGICAL PROBLEMS

(Arranged by the Biometrics Section with the cooperation of the American Association of Economic Entomologists, Eastern Branch)

Chairman: Roger B. Friend, Connecticut Agricultural Experiment Station

Special Plot Techniques Used in Investigations of Fruit Insects
F. Z. Hartzell, New York State Agricultural Experiment Station
Informal presentation of problems in applied entomology for joint discussion by statisticians and entomologists

—8:30 P.M.—

PRESIDENTIAL ADDRESSES

(With the American Economic Association)

Chairman: Frederick C. Mills, Columbia University

The Conditions of Expansion
Sumner H. Slichter, Harvard University, President of the American Economic Association
Government and the Statistician
Winfield W. Riefler, Institute for Advanced Study, President of the American Statistical Association

Tuesday, December 30, 1941

—8:00 A.M.—

BREAKFAST MEETING—OFFICERS AND DIRECTORS

—10:00 A.M.—

DIFFERENTIAL PRICING AND THE GENERAL WELFARE
(With the American Marketing Association)

Chairman: Edwin G. Nourse, Brookings Institution

Types of Differential Pricing
Corwin Edwards, Department of Justice

Examples in Various Industries:

Public Utilities
Burton N. Behling, Interstate Commerce Commission

Milk
E. E. Vial, Surplus Marketing Administration

Agricultural Surplus Programs
A. C. Hoffman, Office of Price Administration

Social Appraisal of Differential Pricing
Morris A. Copeland, Office of Production Management

Discussion: Fritz Machlup, University of Buffalo

THE USE OF 1940 CENSUS DATA BY FINANCIAL INSTITUTIONS

Chairman: Arthur E. Weimer, School of Business, Indiana University

The Use by Financial Institutions of Data from the Population and Housing Census of 1940
David L. Wickens, Consulting Economist

Discussion: Howard G. Brunsman, Bureau of the Census
Joseph Demmery, University of Washington

The Use by Financial Institutions of Data from the Agricultural Census of 1940
Donald C. Horton, Department of Agriculture

Discussion: A. G. Brown, American Bankers Association
Ralph C. Limber, Farm Mortgage Conference of Life Insurance Companies
Norman J. Wall, Department of Agriculture

ANIMAL GROWTH

(Arranged by the Biometrics Section)

Chairman: Antonio Ciocco, U. S. Public Health Service

Determination of the Age of Deer by Means of the Discriminant Function
Besse B. Day and Marion M. Sandomire, Department of Agriculture

Effect of Protein-Mineral-Vitamin Supplements to Barley upon the Growth Rate of Hogs
D. B. DeLury, University of Toronto, and E. W. Crampton, Macdonald College

Analysis of the Relative Growth of Organisms
O. W. Richards and A. J. Kavanagh, Spencer Lens Company

New Formulations in the Study of Rat Growth
T. F. Zucker and Lois Zucker, Columbia University College of Physicians and Surgeons

Tuesday, December 30, 1941

—10:00 A.M.—

RECENT ADVANCES IN MATHEMATICAL STATISTICS

(With the Institute of Mathematical Statistics and the Econometric Society)

Chairman: Harold Hotelling, Columbia University

Recent Advances in Mathematical Statistics

I. Burton H. Camp, Wesleyan University

II. Cecil C. Craig, University of Michigan

HOW SHALL AMERICAN LABOR'S STANDARDS AND SECURITY BE SAFEGUARDED NOW AND DURING THE POST-WAR EMERGENCY?

(With the American Association for Labor Legislation)

Round Table

Chairman: J. Douglas Brown, Office of Production Management

Discussion: Arthur J. Altmeyer, Social Security Board
Ewan Clague, Social Security Board
Herman Feldman, College of the City of New York
Clinton S. Golden, Steel Workers Organizing Committee
Marion Hedges, International Brotherhood of Electrical Workers
Frank Herring, National Resources Planning Board
Lewis L. Lorwin, National Resources Planning Board
W. H. Winans, Union Carbide and Carbon Company

—12:30 P.M.—

LUNCHEON MEETING—THE ECONOMIC OUTLOOK, 1942

Chairman: Winfield W. Riefler, President of the American Statistical Association

The Task of Industry in 1942

Donaldson Brown, General Motors Corporation

The Prospect for 1942

W. Randolph Burgess, National City Bank of New York

—2:30 P.M.—

EXPERIMENTAL DESIGN IN RELATION TO SAMPLING

(Biometrics Section with the Institute of Mathematical Statistics)
Boyce Thompson Institute, Yonkers, N. Y.

Chairman: W. J. Youden, Boyce Thompson Institute

Biological Interpretation of Interactions

W. C. Jacobs, Cornell University

Adapting the Design to the Experiment

G. M. Cox, North Carolina State College

Sampling Theory When the Sampling Units Are of Unequal Size

W. G. Cochran, Iowa State College

Sampling Errors of Systematic and Random Surveys of Cover Type Areas

J. G. Osborne, Forest Service

Various experimental designs on display in the greenhouses of the Institute before and after the session

Minutes of the Annual Business Meeting

The American Statistical Association convened for the 103rd annual business meeting at 9:00 A.M., Monday, December 29, 1941, at the Biltmore Hotel in New York City. President Winfield W. Riefler presided.

The minutes of the 102nd annual business meeting were approved as published in the JOURNAL.[1]

The Secretary read the Annual Report of the Board of Directors.[2]

President Riefler announced the appointment of Lowell J. Reed to the Committee on Fellows for the period ending at the annual meeting in 1946. The report of the Committee on Fellows was read by its Chairman, F. Leslie Hayford. The Committee announced the election of the following Fellows: J. Frederic Dewhurst, George O. May, Vergil D. Reed, Bradford B. Smith, H. R. Tolley, and P. K. Whelpton.

The report of the Nominating Committee was presented by the Secretary. The Secretary reported that he had received no nominations by petition. It was voted to instruct the Secretary to cast one ballot for the candidates presented by the Nominating Committee. The ballot was cast and the following officers and directors were elected:

President: ALFRED J. LOTKA, Metropolitan Life Insurance Company

Vice-Presidents:

Collection and Classification of Data, and Administration of Statistical Agencies
CALVERT L. DEDRICK, Bureau of the Census

Statistical and Actuarial Methods and Technique, and the Teaching of Statistics
WALTER BARTKY, University of Chicago

Facts and Methods Pertaining to Sociology, Social Welfare Problems, and Labor Statistics
HOWARD B. MYERS, Work Projects Administration

Facts and Methods Related to Biometry, Vital Statistics, Psychology, and Education
FRANK W. NOTESTEIN, Office of Population Research, Princeton University

Facts and Methods Bearing upon Economics and Economic Theory
NORMAN J. SILBERLING, Stanford University

Facts and Methods Pertaining Primarily to Business
HENRY B. ARTHUR, Swift & Company

Facts and Methods Pertaining to Financial Institutions
HAROLD V. ROELSE, Federal Reserve Bank of New York

Facts and Methods Pertaining to Marketing
ALFRED N. WATSON, Curtis Publishing Company

Directors: (For the terms expiring at the Annual Meeting in 1944)
WINFIELD W. RIEFLER, Institute for Advanced Study
WALTER A. SHEWHART, Bell Telephone Laboratories

[1] This JOURNAL, 36 (March, 1941) 131.
[2] This JOURNAL, pp. 131–35.

Secretary-Treasurer: RICHARD L. FUNKHOUSER, American Statistical Association

The President presented the recommendations of the Board of Directors for amendments to the Constitution. During the discussion one modification was suggested in the proposal offered by the Board of Directors. It was voted that Articles V, VI, VII and VIII of the Constitution be amended to read as follows:

Article V—Officers

The officers of the Association shall be a Chairman of the Board, a President, a number of Vice-Presidents, a Secretary, a Treasurer, and an Editor. The duties of the officers shall be those usually devolving upon such officers except as may be otherwise provided by the By-Laws.

Each of the Vice-Presidents shall represent a distinct field of statistical interest. The number and fields of interest of the several Vice-Presidential offices shall be fixed in accord with the provisions of the By-Laws. The Vice-Presidents shall assist and advise in the preparation for annual and special meetings and in other matters concerning the interests of the groups they respectively represent in the membership of the Association. In case of the incapacity of the Chairman of the Board to perform the duties of his office on any occasion the President shall act as Chairman of the Board.

Article VI—Board of Directors

There shall be a Board of Directors for the government of the Association consisting of the Chairman of the Board, the President, the Secretary and six elected Directors. The Secretary shall call a meeting of the Board of Directors on the request of the Chairman of the Board, the President, or any three Directors and shall notify the Directors of such meetings at least seven days in advance. Three Directors shall constitute a quorum at any meeting regularly convened.

Article VII—Elections

The officers of the Association other than the Chairman of the Board and the Editor shall be elected each year at the Annual Meeting and shall serve until the close of the Annual Meeting at which their successors are elected. Upon the expiration of his term of office, the President shall become the Chairman of the Board for a term of one year. The Editor shall be appointed by the Board of Directors. The six elected Directors shall be elected for terms of three years each, two of which terms shall expire at the close of each Annual Meeting. No elected Director may serve two terms in immediate succession. The Board of Directors shall fill vacancies in any office and in its own membership until the next Annual Meeting.

The President shall appoint a Nominating Committee of three members, the personnel of which shall be announced to the membership at least three months before the Annual Meeting. At least thirty days before the Annual Meeting the Nominating Committee shall report to the membership its nominations of officers and Directors to be elected at the next Annual Meeting.

Twelve or more Regular Members or Fellows may make nominations for any office by submitting a signed petition to the Secretary of the Association not less than twenty-four hours before the beginning of the business meeting at which the election is scheduled to take place. In case a nomination by petition is made, the Nominating Committee may present to the Annual Meeting the names of one or more additional nominees for the same office.

When but one person has been nominated for a given office, that person may be elected at the Annual Meeting by a viva voce vote. If there are two or more nominations for the same office, it shall be the duty of the Secretary to announce such nominations. In the latter case, the election for this office shall be by secret ballot, and the Chairman of the Board shall appoint three judges of election who shall count the ballots. The judges of election shall decide any questions that may arise as to the validity of any ballot, and shall declare the results of

the election. The candidate who receives a plurality of the votes cast for that office shall be declared elected. The Chairman of the Board shall vote only in case of a tie vote, in which event he shall cast the decisive ballot.

Article VIII—Meetings

Annual and occasional meetings of the Association shall be held at such times and places as the Board of Directors may designate. The Chairman of the Board shall have charge of the general program of such meetings with the assistance and advice of the Vice-Presidents, the Secretary, and the Section Committees.

Theodore H. Brown proposed two amendments to the By-Laws. It was voted to amend By-Law 4 to read as follows: "The field or fields of interest to be represented by each Vice-President shall be specified by the Board of Directors. The Nominating Committee shall make a nomination for each field of interest except that with the approval of the Board of Directors the Committee may nominate not more than eight Vice-Presidents by assigning to some of them responsibility for representing more than one field of interest." It was voted to amend By-Law 9 to read as follows: "These By-Laws may be amended at any Annual Meeting of the Association by a two-thirds vote of the members present and voting, or they may be amended through a mail ballot by a three-fourths vote of the members replying and voting."

The Secretary announced that the Annual Meeting of the Association in 1942 would probably be held in Cleveland, Ohio, and that the Annual Meeting in 1943 would probably be held in Chicago, Illinois.

The Secretary announced that the Association's headquarters office had been moved to the American University, Nebraska and Massachusetts Avenues, N.W., Washington, D. C.

The meeting was adjourned.

R. L. FUNKHOUSER, *Secretary*

Report of the Board of Directors

During the one hundred and two years since the founding of the American Statistical Association, our society has seen the Nation go through several periods of major warfare. As the current year draws to a close we observe the country embarking upon another armed conflict. We know little of what the years ahead may hold in store, but it is evident that our Association, like other institutions throughout the country and like the Nation itself, faces difficult times calling for courage and foresight.

During the first World War enormous stimulus was given to the development of more adequate and more reliable statistical information. The resulting impetus brought about the establishment of many of the existing current measurements of our social and economic life and laid the basis for more timely information than had been available hitherto. At present, as the Nation concentrates its energies and skills on the task of winning the War, we witness a further expansion of our statistical resources. Perhaps there has never been a time in which we are, as a Nation, more dependent upon reliable and timely information as a basis for sound policy and efficient administration.

During this period of transition from a defense to a war-time economy, the Association has maintained the essential functions that marked its program in the past, expanded its resources, and modified its program to meet changing conditions. The following review of the year's work bears witness to these developments.

Study of Defense Statistics: Your officers recognized the immense task encountered by the defense agencies in meeting the informational needs of the preparedness program. The Board of Directors requested the President to study the organization of statistical work of the Federal Government in an effort to appraise the problems of duplication, confusion, and inefficiency resulting from the expansion in the requirements for statistics by the defense agencies. Suggestions for the improvement of the statistical work relating to the defense program were offered informally on several occasions. Through its officers the Association is represented on a number of committees studying the statistical problems in more detail and preparing recommendations for their solution.

Encouragement is also being given to improvements in statistics in two realms closely connected to the defense program. One has to do with the existing information on the qualifications and skills of those composing the statistical profession. It is recognized that the efficiency and reliability of our statistics depend in large measure upon the effective use of competent statistical personnel. The second relates to the development of statistical indexes in more convenient and more comparable terms for the use of consumers of statistics. We have studied the possibilities of greater uniformity in the base periods of general-purpose indexes and surveyed the problems encountered in adopting a uniform base for their presentation.

Changes in Organization: During the year the Board of Directors and the officers continued to study the problems of organization of the Association and consulted with the Fellows and recent officers to get their reactions and suggestions. After careful deliberation it was decided to propose to the membership at this annual meeting that we make a number of changes in our organization in order to distribute more efficiently the responsibilities devolving upon the officers and to strengthen the preparation of the program for the annual meetings. The Directors also adopted a broad set of policies for their own guidance designed to stimulate the Association's Chapters and to encourage the organization of Sections for important subject-matter fields. We believe that it is important to keep the organization of the Association flexible in order to meet the changing conditions and to experiment with promising solutions to specific problems as they arise.

Committees: Committees of the Association have continued to render useful service to statisticians in several fields. The Census Advisory Committee held four meetings to review the program of the Sixteenth Decennial Census and to encourage greater emphasis on the analysis and interpretation of census data. The Committee considered the effect of the defense program on the activities of the Bureau and noted the resources for statistical services

that can be made available to defense agencies. It reviewed the proposals for a quinquennial census of industry and business and the collection of current business statistics; it discussed the proposals for a quinquennial census of population and housing combined with sample censuses in intervening years. The Committee advised in connection with the urgency of securing the funds necessary to bring the Sixteenth Decennial Census to a successful conclusion.

The Joint Committee on Relief Statistics held several meetings and rendered advice on a number of problems in the field of relief and welfare statistics. The Committee decided to resume the preparation of the *Bulletin of Information of Relief Statisticians* and arranged sessions at the annual meetings of the sponsoring societies. The Joint Committee on Occupational Classification met to develop policies on the tabulation of the occupational and industrial statistics of the Census of 1940. During the year three intermediate classification lists were prepared to be applied to areas where the complete detail could not be used; each is in harmony with the Convertibility List prepared by the Committee several years ago. A subcommittee is at work on the classification of new or changing occupational groups, with particular reference to titles peculiar to defense activities. The Committee on Census Enumeration Areas assisted in the plans for tabulations of census tract data and gave leadership to their wider use by statisticians. Members of the Joint Committee for the Development of Statistical Applications in Engineering and Manufacturing cooperated actively with many groups concerned with the application of statistical methods to defense problems in the field of industry. They also assisted in arranging scientific meetings and in the presentation of defense courses in statistical methods. The Joint Committee on Metropolitan Districts completed its task of recommending the Metropolitan Districts for a large number of densely populated areas and these were established as a part of the regular Sixteenth Decennial Census procedure. The Committee on Sampling was established during the year and promises to make an important contribution to the sound use of sample techniques in both private and governmental statistics.

Chapters: The Chapters of the Association maintained increasingly worthwhile activities in meeting the needs and interests of members in their respective areas. The Albany Chapter was responsible for the preparation of a printed index to existing information in the New York state government. A similar guide was prepared by the Harrisburg Chapter covering available data in Pennsylvania. Several of the Chapters are engaged in compiling directories of statisticians in their localities. These and similar enterprises supplemented the series of meetings arranged by most of the Association's Chapters. The Committee on Chapters has assisted in suggesting out-of-town speakers for a number of these meetings and has exerted leadership in the establishment of new Chapters. Chapters were launched in Montgomery, Alabama, and at the University of Illinois. There is evidence of growing interest in a dozen other centers in the establishment of local Chapters.

Biometrics Section: The Biometrics Section experienced an especially

fruitful year. Its membership, formerly composed almost entirely of professional statisticians, was doubled by enrolling biologists and similar specialists employing statistical techniques in their research. It arranged a number of scientific meetings with biological groups, in addition to the sessions with our Association and those organized by a special committee on the West Coast. Plans are under consideration for the near future calling for additional scientific meetings with biological societies, improvements in the organization of the Section, and other developments that will expand the usefulness of the Section's activities.

Publications: Four regular numbers of the JOURNAL were printed during the year. The first six issues of the expanded *Bulletin* were prepared. The Statistical News and Notes and the Chapter Activities, hitherto published in the JOURNAL, were transferred to the more frequent current issues of the *Bulletin*. The *Index to the Journal of the American Statistical Association*, covering the period from 1888 through 1939 and prepared last year, appeared early in the current year. A large number of copies were purchased by members and subscribers. The Association is grateful to the Rockefeller Foundation for making possible this important contribution to the usefulness of the JOURNAL. By action of the Board of Directors the net proceeds from the sale of the *Index* are being added to the Centenary Sustaining Fund.

Membership and Finances: Under authority of the amendment to By-Law 2 adopted last year, the Board of Directors established a special rate of dues of $3.00 for new members under 30 years of age during the first year of membership. In response to this action and the conscientious efforts of the Committee on Chapters and several local committees about 350 young statisticians joined the Association at the reduced rate. In addition nearly 275 statisticians and others interested in the Association's work became members at the regular rate of dues, bringing the total number of new members to more than 600. This produced a net increase in the membership during the year of about 350, so that at the end of the year our total membership is nearly 3,100.[1]

Largely as a result of this encouraging growth in membership our income from regular sources during 1941 was approximately $1,300 larger than the corresponding income a year ago. In addition, the sum of $2,500 was transferred from the Centenary Sustaining Fund to meet current operating expenses for the year.

It is difficult to forecast the effect of war conditions on the financial position of the Association during the period immediately ahead. However, your officers believe that with the further growth of statistical activity connected with the war effort we may expect a sustained growth in the Association's membership. We also anticipate increases in the expenses involved in maintaining our program. Some evidences of these increases in the expenses appeared during 1941 and if they become more severe and more rapid they may

[1] These estimates were superseded by figures compiled after the close of the year. See the Report of the Secretary, this JOURNAL, pp. 135-36.

require us to make some modifications in our plans. In order to effect economies that will not impinge seriously on our functions it was decided to move the Association's office to a new location furnished to us without expense through the generosity of the American University in Washington.

Responsibilities for the Future: In a period like the present there is an especial need for trustworthy and prompt statistical information, efficiently organized and accurately interpreted. The requirements relate not only to the defense industries but also the broader phases of our complex social and economic structure and its operation under rapidly changing conditions.

Consequently, while we must make adjustments in our Association's activities and plans as a result of the national emergency we need constantly to be on the alert to maintain a balanced program and to encourage lasting improvements in statistics and statistical methods. Moreover, as a professional society we should take part in preparing for the problems that will face the Nation in the period marked by a return to peace-time activities. These problems and the information that will be needed to cope with them will be no less urgent than those we face during the armed conflict. Our responsibilities extend beyond the present and the immediate future and require that we keep alive the realization that the Nation's task of winning the peace is fully as important as that of winning the War.

W. LEONARD CRUM	WINFIELD W. RIEFLER
R. L. FUNKHOUSER	FREDERICK F. STEPHAN
F. LESLIE HAYFORD	O. C. STINE
LOWELL J. REED	WILLARD L. THORP

Report of the Secretary

Preliminary membership statistics for the calendar year 1941 are included in the *Report of the Board of Directors.* On the basis of more recent information, however, it is possible to summarize the final statistics about the membership of the Association. During 1941 the total membership increased from 2,734 to 3,181, reflecting a net growth of 447 members. During the year 705 new members were elected and 20 former members were reinstated. Twenty members were lost by death during the year, including two Honorary Members, one Fellow, and one Life Member; 103 members resigned; the memberships of 155 were removed from the active rolls for failure to pay dues.

Membership statement, December 31, 1941

Honorary members	14
Corporate members	4
Fellows	102
Regular members	3,061
Total membership	3,181

One Fellow was a Contributing Member during 1941. There were no new Life Members during the year but the death of one Life Member was reported reducing the total number of Life Members to 36.

The death of the following members was recorded during the year: The Rt. Hon. Lord Stamp of Shortlands, and Gustavus Thirring, *Honorary Members;* James W. Glover, *Fellow;* Carl L. Alsberg, M. Ada Beney, Robert G. Brown, Arthur R. Bruder, R. L. Crandall, Robert F. Foerster, Bernadino A. Giliberty, Hubert W. Horn, Olin Ingraham, John Wells Morss, S. B. Pearmain, Paul A. Ryan, C. E. Sniffen, Hazel I. Spicer, M. W. Thompson, Jessamine S. Whitney, and Evelyn L. Yeomans, *Regular Members.*

R. L. FUNKHOUSER, *Secretary*

Report of the Treasurer

During the calendar year 1941 the Association's income from regular sources exceeded that in 1940 by about $1,600. In addition, the gross proceeds from the sale of the *Index to the Journal* amounted to nearly $600, bringing the total of the increase over 1940 to almost $2,200. An expansion of nearly $900 was recorded in membership dues; income from subscriptions increased about $300, and advertising revenue was approximately $400 greater than in 1940. In order to supplement the income from regular sources, the Board of Directors authorized an appropriation of $2,500 from the Centenary Sustaining Fund, compared with $1,500 during the preceding year. This expanded transfer to the general account of the Association largely offset the termination of the Rockefeller Foundation grant for general purposes.

The regular program of the Association's activities during 1941 resulted in an increase in expense of a little more than $600 over the corresponding total in 1940. Both office expenses and the cost of the *Bulletin* were larger than the year before, but salaries and wages and the cost of the JOURNAL were reduced. The total expense for the year was nearly $1,400 lower than in 1940.

The net excess of income over expenses, amounting to nearly $400, was accounted for largely by the net proceeds from the sale of the *Index to the Journal*. By action of the Board of Directors, the net proceeds from the sale were transferred to the Centenary Sustaining Fund.

The balance sheet and other detailed financial statements, with comparative figures for 1940, are included in the Report of the Auditors.

R. L. FUNKHOUSER, *Treasurer*

Report of the Auditors

To the Board of Directors of
American Statistical Association

We have examined the balance sheet of the American Statistical Association as at December 31, 1941 and the statements of income and surplus for the year then ended, have reviewed the accounting procedures of the Association and have examined or tested accounting records and other supporting evidence, by methods and to the extent we deemed appropriate.

The recorded cash receipts for the year were traced to the deposits shown on the bank statements and the amounts for dues and subscriptions were tested with the membership and subscription records. The paid checks and relative vouchers were inspected in support of the cash disbursements for the year. The cash balances and the securities owned as at December 31, 1941 were confirmed by inspection or by certificates obtained direct from the depositaries. We did not check the membership and subscription records in detail or make any independent verification of the inventory of old journals, the office records of which are based in part on data assembled in prior years, no recent physical inventory having been taken.

In accordance with a resolution of the Board of Directors, March 31, 1936, the life membership reserve is computed on the basis of the combined annuity table of mortality with assumed interest at 4% per annum and an assumed annuity of $5.00 per life member. The amount treated as income in each year represents the excess of the reserve at the beginning of the year plus interest for the year and new life memberships receipts over the required reserve at the end of the year.

In our opinion, the accompanying balance sheet and related statements of income and surplus present fairly the position of the American Statistical Association at December 31, 1941 and the results of its operations for the year, in conformity with generally accepted accounting principles applied on a basis consistent with that of the preceding year.

PRICE, WATERHOUSE & CO.

March 9, 1942

AMERICAN STATISTICAL ASSOCIATION

BALANCE SHEETS

Assets		December 31, 1941	December 31, 1940
Cash in bank and on hand		$ 1,753.60	$ 2,399.36
Accounts receivable		325.01	156.83
Investments:			
United States Savings Bonds, at redemption value		5,186.08	5,052.00
Stocks, at cost (at market quotations $3,875 and $4,158, respectively)		5,793.50	5,793.50
Inventory of old Journals, at approximate cost applied to salable quantities		1,770.91	1,791.75
Furniture and equipment, at cost less depreciation		541.81	627.07
Deferred charges		150.31	
		$15,530.83	$15,820.51
Liabilities			
Accounts payable		$ 1,183.18	$ 245.91
Special account—Allied Social Science Associations		481.81	308.33
The Rockefeller Foundation grant for the Index to the Journal:			
Total amount	$5,000.00		
Less—Portion expended in 1939 and 1940	3,627.69		
	$1,372.31		1,372.31
Less—Portion expended in 1941	912.74		
Amount returned to the Rockefeller Foundation	$ 459.57		
Centenary Sustaining Fund, per statement		3,018.45	4,450.47
Life membership reserve		2,315.70	2,411.01
Unearned income:			
Dues		1,207.12	910.42
Subscriptions		1,600.76	1,363.02
Surplus, per statement		4,073.72	4,007.24
		$15,530.83	$15,820.51

AMERICAN STATISTICAL ASSOCIATION

INCOME STATEMENTS

	Year ending December 31, 1941	1940
INCOME:		
Dues—current year	$14,092.00	$13,161.71
Dues—prior years	31.25	37.50
Life memberships	192.60	280.03
Subscriptions	3,723.25	3,304.50
Advertising	832.62	458.04
Reprints	174.40	180.10
Journal sales	203.07	284.33
Index to the Journal sales	506.00	
Special publications	71.07	18.04
Miscellaneous	128.25	91.18
Dividends and interest (after deducting $90.07 in 1940 and $90.48 in 1941 apportioned to life membership reserve)	208.34	191.98
	$20,222.85	$18,004.91
Appropriation from Centenary Sustaining Fund	2,500.00	1,500.00
The Rockefeller Foundation grants:		
For general purposes		1,500.00
For Index to the Journal (appropriation to cover expenses for year, per contra)	912.74	2,590.98
	$23,635.59	$23,655.89
EXPENSES:		
Journal—printing, mailing and reprints	$ 4,722.89	$ 5,048.25
Bulletin	1,441.75	402.84
Salaries and wages	11,082.20	12,200.20
Unemployment compensation tax	238.06	281.30
Rent	655.00	900.00
Office supplies, printing and mimeographing	932.70	725.92
General postage and carriage	710.04	535.20
Telephone and telegraph	155.62	212.45
Travel expense—officers	387.20	370.07
Mimeographing—committees	59.46	19.92
Storage of old Journals	72.00	72.00
Cost of old Journals sold	59.64	40.70
Miscellaneous expense	475.35	200.54
Depreciation of furniture and equipment	95.61	110.00
Membership campaign	170.80	
Membership directory		701.15
Expense—Index to the Journal:		
Expended from Rockefeller Foundation grant, per contra	912.74	2,590.98
Expended from Association funds	211.35	
	$23,274.46	$24,030.88
Balance carried to surplus	$ 361.13	$ *983.90*

AMERICAN STATISTICAL ASSOCIATION

SURPLUS STATEMENT

	Year ending December 31, 1941	1940
Balance at beginning of year	$4,967.24	$5,951.23
Add—Excess of income over expenses for the year, per income statement	361.13	
	$5,328.37	$5,951.23
Deduct—Excess of expenses over income for the year, per income statement		$ 983.99
Transfer to the Centenary Sustaining Fund of net proceeds from sale of the Index to the Journal	$ 354.65	
	$ 354.65	$ 983.99
Balance at end of year	$4,973.72	$4,967.24

STATEMENT OF CENTENARY SUSTAINING FUND

	Total	Year ending December 31, 1941	Prior to 1941
Contributions and pledges (relating to the support of the activities of the Association for a period of five years from January 1, 1940 to December 31, 1944)	$10,696.02		
Less—Pledges not collected at December 31, 1941, including $2,700.40 not due at that date	2,701.25		
Amounts received from contributors	$ 7,994.77	$1,202.00	$6,792.77
Interest received on bank savings account and increase in redemption value of securities	135.56	111.33	24.23
Transfer to Centenary Sustaining Fund of net proceeds from the sale of the Index to the Journal	354.65	354.65	
Total cash receipts	$ 8,394.98	$1,667.98	$6,727.00
Less—Expenses of campaign (printing, postage and temporary assistance)	776.53		776.53
Net receipts from campaign	$ 7,618.45	$1,667.98	$5,950.47
Appropriations by the Board of Directors to the general account of the Association in accordance with the budget: Year 1940 $1,500.00; Year 1941 2,500.00	4,000.00		
Balance December 31, 1941, per balance sheet	$ 3,618.45		

List of Committees and Representatives for 1941

Committee on Fellows[1]

F. Leslie Hayford
Frank A. Ross
W. Leonard Crum
John Rice Miner
Theodore H. Brown

Committee on Nominations

Theodore O. Yntema, *Chairman*
Halbert L. Dunn
Dorothy S. Thomas

Committee on Investments[2]

F. Leslie Hayford, *Chairman*
George O. May
Walter W. Stewart

Biometrics Section Committee

C. I. Bliss, *Chairman*
Joseph Berkson, *Secretary*
A. E. Brandt
Alfred J. Lotka
Hugo Muench, Jr.
J. Neyman
Carroll E. Palmer
George W. Snedecor
Douglas E. Scates

Census Advisory Committee

William F. Ogburn, *Chairman*
Murray R. Benedict
Paul T. Cherington
J. Frederic Dewhurst
Luther Gulick
Frederick F. Stephan
Willard L. Thorp

Committee on Census Enumeration Areas

Howard Whipple Green, *Chairman*
Clarence E. Batschelet
Ernest M. Fisher
Charles S. Newcomb
Vergil D. Reed
Leon E. Truesdell
Shirley K. Hart
Lent D. Upson

Joint Committee on Metropolitan Districts (With the American Marketing Association and the Chamber of Commerce of the United States)

Paul T. Cherington[3]
T. W. Howard[4]
Glenn E. McLaughlin

[1] With terms expiring at the end of 1941, 1942, 1943, 1944, and 1945, respectively.
[2] Appointed in accordance with By-Law 5 to supervise the investment of the Association's surplus funds.
[3] Appointed by the American Marketing Association.
[4] Appointed by the Chamber of Commerce of the United States.

Joint Committee on Occupational Classification (With the Division of Statistical Standards, Bureau of the Budget)

Howard B. Myers, *Chairman*
Gladys L. Palmer, *Secretary*
E. Dana Durand
Meredith B. Givens
Edward D. Hollander
Merrill G. Murray
Leon E. Truesdell
Emmett H. Welch

Joint Committee on Relief Statistics (With the American Public Welfare Association)

Ralph G. Hurlin, *Chairman*
Paul Webbink, *Secretary*
Neva R. Deardorff
Emil Frankel
Anne E. Geddes
Helen R. Jeter
Howard B. Myers
Emerson Ross
Saya S. Schwartz
Herman M. Somers
Emmett H. Welch

Committee on Statistics of Hospitals

Graham L. Davis, *Chairman*
Joseph Berkson
Selwyn D. Collins
Edwin Crosby
Neil A. Dayton
Halbert L. Dunn
Gerhard Hartman
M. R. Kneifl
R. G. Leland
George St. J. Perrott
C. Rufus Rorem
Charles G. Roswell
A. Hardisty Sellers

Committee on Chapters

Robert R. Nathan, *Chairman*
Henry B. Arthur
Edwin B. George
Lester S. Kellogg
William A. Sturm
Charles W. Vickery, Jr.
Robert R. Williams

Committee on Sampling

Frederick F. Stephan, *Chairman*
George W. Snedecor
Alfred N. Watson
Samuel S. Wilks
Theodore O. Yntema

Advisory Committee to the National Roster of Scientific and Specialized Personnel

Halbert L. Dunn
Walter Mitchell, Jr.
Frederick F. Stephan
Ralph W. Tyler
Samuel S. Wilks

Representative on the Board of Directors of the National Bureau of Economic Research

Winfield W. Riefler

Members of the Social Science Research Council[1]

O. C. Stine
Stuart A. Rice
Edwin B. Wilson

Representative on the Council of the American Association for the Advancement of Science

Walter A. Shewhart

Representatives on the Joint Committee for the Development of Statistical Applications in Engineering and Manufacturing

Walter A. Shewhart
Churchill Eisenhart

Representative on the Sectional Committee on Standards for Graphic Presentation

A. H. Richardson

Member of the American Documentation Institute

Frederick F. Stephan

Representative on the Advisory Board of the American Year Book

Frank W. Notestein

Representative on the Advisory Committee to the Census Library Project

Richard O. Lang

[1] With terms expiring at the end of 1941, 1942, and 1943, respectively.

BOOK REVIEWS

GLENN E. MCLAUGHLIN
Review Editor

Introduction to Medical Biometry and Statistics, by Raymond Pearl. Philadelphia: W. B. Saunders Company. 1940. xv, 537 pp. $7.00.

A few months after publication of this book and after I had accepted an invitation to review it, the author, Raymond Pearl, died. I regarded Pearl as my most influential teacher as well as a good friend. It is natural, therefore, that the task of reviewing his last edition of a book which was the foundation of my own statistical meanderings should invite a somewhat different treatment from what ordinarily is due another edition of a well known text.

Pearl never did anything in a way that anyone else would do it, and the trait is exemplified in his *Medical Biometry and Statistics*. It is a good solid elementary text—the best in its field twenty years after its first issue—but it is more; it is Pearl's personality. In it the scholar, the prolific reader, the wit, the English stylist, are on exhibit, as well as the pontiff and the paradoxer. In his career Pearl published an immense amount of work and minuteness of statistical detail was characteristic of it, but always it was enlightened by a well defined general scientific *Anschauung*. His great scientific hero, and I suspect his conscious model, was Galton and he exhibited the same versatile naturalist's interest in an enormous variety of subjects, and the same desire to comprehend them quantitatively through the instrumentality of statistics. But, as revealed in the incisive introductory chapter of this book, his fundamental view was that the statistical method is only ancillary to the experimental. Characteristically one can find in his work a parallel series of investigations, on the one hand strictly statistical, on the other experimental. Illustrating this is his work on the application of the logistic curve to population growth, in which there were extensive investigations, with Lowell J. Reed, on human populations and at the same time intensive experiments on Drosophila. Similarly his idea that longevity is related inversely to the rate of expenditure of energy, first propounded with the support of occupational mortality statistics, was followed by elaborate experiments with the fruit fly and Cucumis melo.

The second chapter of the book deals with land marks in the history of biostatistics. While the author himself disparages the dry method of tabular historical outline which is the basis of the chapter, the selections for emphasis, the quotations, the comments, as well as the illustrations, furnish an entertaining as well as informative *aperçu* of the high points in the development of biomedical statistics. Chapters III–VI deal with the raw data of biometrics, their selection, and tabular and graphic presentation.

Beginning with Chapter VII the analytical technic of statistical methods is dealt with. Here students of statistics who have entered the field in recent years, and especially those who have been influenced by the ideas of R. A.

Fisher, will find an unfamiliar approach, and even what they may consider to be erroneous. Pearl was well aware of relatively recent "revolutionary" changes in statistics, and the failure to modify his text extensively in accord with them, wise or not, was deliberate. There is in the book no "theory of small samples," no "analysis of variance," no great emphasis on the "testing of hypotheses," nor any great show of the new statistical vocabulary currently in the making. These omissions are, in my opinion, in the main sound.

The eighth chapter, on life tables, is the best brief presentation of the subject of which I am aware, and in some respects the best regardless of length. The title of the book would seem to limit us to medical biometry, but this chapter exhibits the Pearlian touch, and the life tables of flies, roaches and automobiles are also dealt with. Here, as elsewhere in the book, crops out that, to Pearl, man was just another organism, though withal an extraordinary one, which a biologist could examine with interest.

In the ninth chapter, on adjusted death rates, Pearl comes near to losing his sense of humor over the existing confusion in terminology. (Or is it that this reviewer lost his?) He says in effect that since the statistical world has not had the sense to adopt his terminology, he will follow the rather inept official recommendations. Then he proceeds *not* to follow them!

The tenth chapter deals with the sampling error concept and presents some new tables for determining whether an observed difference between means is significant. The eleventh chapter on the theory of probability deals with the binomial expansion, its approach to the normal curve, and the estimation of the variability of a rate, while the following chapter presents the chi-square test and the four-fold table. In Chapters XIII and XIV on measurement of variation are taken up the calculation of moments and the derived statistics, the mean, median, mode, standard deviation, skewness and kurtosis, as well as some other matters such as class limits, the retention of decimal places in published constants, the comparison of groups with regard to composite items and graphic presentation of variability on the basis of range. These chapters are not happily headed so far as indicating their contents is concerned, and the soundness of some of the new suggestions contained in them is debatable. Chapters XV, XVI, and XVII deal with correlation and simple curve fitting.

The last chapter, XVIII, deals exclusively with the logistic curve, and here the guiding hand of Professor Lowell J. Reed can be discerned. It contains a clearly presented example of the fit of the curve by least squares using successive approximations to a Taylor expansion, and a bibliography on the logistic curve which we may consider definitive. Thus the last book of Pearl ends with a subject for which Pearl is perhaps most widely known, the logistic curve and population growth.

The development of Pearl's interests from the simple flora and fauna that furnish the *matériel* of the professor of biology to the "higher" forms and finally to man was continuous. By the time of his association with the Johns Hopkins University, where I knew him, his central interest had crystallized in *homo sapiens*. To study man in all his complexities sympathetically, and

yet to regard him always as an animal for cold scrutiny was his objective. He held consciously to the broad purpose of applying general biologic principles to the study of man and coined the special phrase "human biology" to convey this. His *Studies in Human Biology* and his later founding of the journal *Human Biology* were expressions of this idea. At the time of his death he was preparing for publication a volume on *Man the Animal*.

Pearl was of the type of the creative scientific worker rather than of the class room teacher. As a class lecturer he left a good deal to be desired, but in a broad sense he was a great teacher, for he affected mature students profoundly. Not a few, coming to his courses incidentally, were seduced by biometry as he represented it to make statistics their life work. It can fairly be said that his influence was greater than that of any single other person in the establishment of the biometric method in this country.

To many, Pearl was associated with a certain basic attitude in biological questions. He believed in the predominant importance of "nature" over "nurture," he emphasized constitutional as opposed to environmental factors in the breakdown of the organism and he laid great stress on genetic factors. He had a "hard boiled" attitude toward ameliorative efforts for improving the environment, when he felt that the basic weakness was in the host organism. He was emphatic in demanding evidence that well intended efforts necessarily produce desirable results, but he made these demands on himself as well as on others. He never uttered an opinion for serious consideration without a heavy support of evidential data amply and clearly tabulated. In the last several years there was notable in his writings what might appear to be a softening in his attitude. For instance he appeared to acknowledge that deliberate efforts at birth control could have a statistically significant effect on population, and that individual medical efforts could have a greater effect on general sanitary conditions than he had been wont earlier to assign to them. I have wondered what was the reason for this apparent change and I believe that it stemmed from the accumulation of data which convinced him. These questions were under continuous investigation in his own laboratory. He had a tremendous respect for the facts, and I believe he was simply convinced by the evidence.

Those who got close to Pearl as graduate students or laboratory colleagues found an amazing personality. He was interested in anything in the realm of the intellect and seemed to have read everything. He had an enthusiasm for Aristotle as though the latter were a contemporary biologist. He possessed a vanity of such colossal proportions as to render it almost impersonal and altogether charming. His weekly journal club meeting was an intellectual frolic where anything from Patten's *Embryology of the Chick* to Hemingway's *Death in the Afternoon* was reviewed with thoroughness and gusto. He taught the search for the objective truth and on all who worked with him he left an indelible mark. He will be remembered by them gratefully and affectionately.

JOSEPH BERKSON

Mayo Clinic

Statistics for Sociologists, by Margaret Jarman Hagood. New York: Reynal and Hitchcock, Inc. 1941. viii, 934 pp. $4.00.

Students of sociology will find in this book an introduction to the statistical tools of their trade rather than an exposition of statistical theories *in vacuo* or in any other medium. But a nine-hundred-page textbook "designed primarily for the first year of statistics for students in sociology" who are presumed to have mastered "no prerequisite of mathematics beyond freshman college mathematics" inevitably raises some troublesome pedagogical dilemmas. Eschewing all higher mathematics while undertaking to explain "all the basic statistical methods which have been used in sociological research and some of the newer ones which have not yet found wide application in sociological fields" has led in this instance to some precarious tightrope walking between rule-of-thumb precepts and such Platonic dialogue as the following:

A. . . . What do other sociologists do about tests of significance in such situations?
B. They do everything imaginable . . .

On the other hand, the text contains many lengthy computations whose rationale is surely beyond the comprehension of most students of freshman college mathematics. Again, although the student is repeatedly and wisely admonished always to consider the basic assumptions of each method that he is tempted to use, his attention is likely to be distracted by occasional digressions from discussion of first principles into minutely detailed instructions for the use of calculating machines to compute various functions on a mass-production basis.

Aside from such shortcomings—probably unavoidable in a text which aims to do so much for novices—the book is not only sound and up-to-date but extremely comprehensive. It should appeal to students who are above the average in intellectual maturity and scientific curiosity. Parts of it, especially those dealing with the newer concepts relating to sampling and analysis of variance and co-variance, should be required reading for many social researchers whose statistical training was derived from the standard textbooks of no more than a decade ago and who do not yet realize that Fisher and others have outlawed inverse probability.

The book is divided into five parts—quantitative methods in sociology, descriptive statistics, inductive statistics, statistics of relationship, and selected techniques for population data. The last part includes treatment of birth and death rates and the construction of life tables. Nine appendix tables embrace integrals and ordinates of the normal curve, *t*, *chi-square*, *F*, *Z*, and the life-table functions *q* and *m*. A systematic collection of the formulae scattered through the text would have enhanced the book's usefulness for reference.

ELBRIDGE SIBLEY

Bowdoin College

The Theory of Econometrics, by Harold T. Davis. Bloomington, Indiana: The Principia Press, Inc. 1941. xiv, 482 pp. $4.00.

The contents of this book are encyclopedic. One finds in it a description of depreciation formulae and of various methods of computing life insurance premiums; the author discusses the concept of utility and presents a detailed theory of international exchange; he devotes considerable space to the problems of taxation and makes passing observations concerning the interrelation between interest rates and the outcome of presidential elections.

In his endeavor to present a complete, up to date picture of the application of quantitative methods of analysis in economics Professor Davis refers to the works of more than three hundred different authors. Numerous bibliographical footnotes are implemented by summaries of selected literature attached to each of the eighteen chapters.

Of the two principal approaches through which quantitative analysis enters the field of economics—pure theory and statistics—both are given equal attention. No attempt, however, is made by Professor Davis to integrate either one in a unified system.

Although essentially a compendium, the book contains here and there original contributions. As many an econometrician with a background of mathematics and physics, the author shows great weakness for analogies between familiar definitions of theoretical physics and some basic concepts of pure economics ("isomorphism between money utility on the one hand and entropy on the other" p. 171 etc.). If his more cautious colleagues should object to Professor Davis' readiness to develop an elegant deductive argument on the basis of a few bald factual and theoretical assumptions, he will be able to point out that some of the best known authorities in the field of business cycles theory reach far bolder conclusions on the basis of much more dubious assumptions—and they get away with it.

The appearance of this volume raises again the question whether "Econometrics" as defined by Professor Davis should and could be usefully treated as a distinct, separate part of economics. It is true that some investigators show a greater facility in and interest for the mathematical formulation of their problems and numerical manipulation of their data than others. The powerful tools of quantitative method have definitely proven their worth in application to economic analysis and thus they are bound to become common property of an ever wider circle of economists. It can be hoped that a working knowledge of the fundamentals of mathematical analysis will soon be as much required from an advanced student of economics as the knowledge of foreign languages. But precisely for this reason a "Theory of Econometrics" can not be more usefully established as a separate field of advanced studies than for example a "Theory of Economics treated with the help of foreign languages." The use of economic problems in teaching mathematics to economists is of course highly advisable, but so is the use of Barone and Böhm Bawerk in teaching them Italian or German. However, before Econometrics have found their natural place in the general body of Econom-

ics a great pedagogical task has to be performed. Professor Davis' book will be of invaluable help to those engaged in teaching mathematics to economists and economics to mathematicians.

WASSILY W. LEONTIEF

Harvard University

The Elements of Statistics, by Elmer B. Mode. New York: Prentice-Hall, Inc. 1941. xvi, 378 pp. $3.50.

Many teachers of statistics must be conscious of the desirability of introducing into their beginning courses more extensive mathematical treatments of the subject than often have been employed in the past. One deterrent has been the lack of suitable textbooks that are at once adequate and still simple enough mathematically to be used with classes in which the students have a mixed (and usually minimum) mathematical training. To many instructors Professor Mode's book may offer a solution. It presupposes no mathematics beyond secondary school geometry and algebra, and yet it is comprehensive and contains much material of value in the study of the elementary mathematics of statistics. Although the author employs an elaborate symbolism it is always carefully defined, and the subject-matter is explained in great detail so that small points which often trouble the beginner are not overlooked. The book is constructed upon the premise that the science of statistics may be studied from a broad general standpoint, and that once the principles are learned, applications may be made to any particular field of interest.

Fundamentals are not neglected in this textbook. The chapter on computation, dealing with limits of accuracy, the elimination of unnecessary calculations, and the use of the slide rule, should prove helpful to beginning students to whom these are often small but vexing problems. There are many worked examples to accompany the discussions of methods, and there are an abundance and a great variety of exercises for solution. Somewhat more attention might have been given to the technique of chart construction in the chapter on that subject; but the charts and graphs of the book itself must be cited as excellent examples of statistical draftsmanship.

The chapters on averages and those on dispersion and moments are very well done, and provide material on the mathematical properties of these measures not usually included in an elementary text. Other chapters of outstanding merit are those on frequency distributions, the frequency curve, and curve fitting. On the other hand, the treatment of index numbers seems rather narrow and limited since it is confined almost entirely to the matter of formulas.

A feature of the book is the introduction at an early stage of the concept of probability and the development of it throughout the text. The student is led gradually and carefully to the final chapters in which the topic is brought to culmination in the discussion of binomial distribution, and in which he is given an introduction to sampling theory. Discussion of measures of reliability is confined principally to large samples. The treatment in rela-

tion to small samples is of necessity somewhat limited, in keeping with the scope of the text. The Chi-square function, however, is given more than passing attention, and a number of examples of the practical application of the Chi-square test help to explain the concept.

In general, the book is very well written, and is marked by a clear, lucid exposition of the mathematical principles involved. It is a book that may require close application from the student, but also one that will repay him for the effort.

Frederick R. Ottman

New York University

Business Statistics, by Martin A. Brumbaugh and Lester S. Kellogg. Chicago: Richard D. Irwin, Inc. 1941. xv, 913 pages. $4.00.

This book is a substantial addition to the growing list of excellent texts prepared primarily for college courses in business and economic statistics. It is designed for the statistical consumer rather than the statistical producer; for the ordinary user rather than the specialist.

The emphasis is on the simpler techniques of collecting, presenting, and analyzing business data and upon affording sound bases for the interpretation of statistical materials. The authors do not introduce mathematics beyond arithmetic and intermediate algebra, including series and the use of logarithms. The majority of business men and economists who have occasion to make use of statistics are not helped by the mysterious symbolism and the involved refinements of the mathematical statistician. The layman in statistics sometimes suspects that the advanced specialists are so enamored of their own abstruse mathematics that they fail to appreciate the importance of, and need for, a more democratic acquaintance with sound statistical techniques.

The authors have devoted considerable space to those aspects of statistical method that are especially pertinent to business and economics. The principles of sampling, the collection and tabulation of data, the use of graphs, and the treatment of time series make up the bulk of the text. Less attention is given to averages and frequency distribution than to trends, index numbers and cycles. The book includes a rich variety of illustrative matter not strictly "business statistics," for among the materials used for explanation or example are those referring to prices, wages, foreign trade, cost of living, the public domain, public health, straw ballots, accidents, marriages, air forces, and postal receipts. The reviewer notes with satisfaction the attention given to the limits of the reliability of the various statistical measures; it would not have been out of place to have included also a caution against some of the "horrible examples" of misleading graphic and pictorial representation which unfortunately abound in popular business and general literature. A helpful innovation in a text on statistics is the inclusion of two chapters on the collection and use of materials from library sources.

Kenneth Duncan

Pomona College

A Significance Test for Time Series, by W. Allen Wallis and Geoffrey H. Moore. New York: National Bureau of Economic Research, Inc. 1941, xii, 59 pp. 50 cents.

The authors have examined some of the characteristics of order that are dependent only on the relative magnitudes of the terms in time series and have derived a pseudo χ^2 test for cyclical fluctuations. Their stated objective is to choose "both a form of the population and the characteristic(s) by which randomness is to be judged . . . entirely without reference to the sample."

Starting with an assumed sequence of N terms, no two of which are identical in magnitude, they derive the exact distribution of lengths of rises and declines (phases or runs), excluding the end or "incomplete" runs, for a theoretical arrangement in which the frequency for each length is the average for the $N!$ possible independent arrangements. Although a correction factor is indicated for the case where the end runs are included, the simple expression

$$\frac{2}{(d+2)!}\{1+(d+1)(N-d)\}$$

for the number of runs of length d or over in such a theoretical distribution is not given. An empirical check of the distribution of "complete" runs is shown for three groups of series ($N=25$, 50, and 75). For this purpose the theoretical frequencies are adjusted only with respect to N, the length of the series and the total frequencies are not identical to the corresponding observed totals. This could, of course, have been avoided by including the end runs and counting the runs (rises and declines) of length, $d=0$. When these are included, the number of runs in any arrangement is $(N+1)$.

The distribution of χ_p^2 is obtained by using a three term comparison ($d=1$, 2, or 3 or over) between the observed and the theoretical distributions of "complete" runs for each sample used in the empirical check. Exact distributions (except for a minor discrepancy in that $(2N-7)/3$ is taken as the total theoretical frequency in each case) are shown for $N=6$ to 12 inclusive and those are compared with the empirical distributions to reach an approximate distribution of χ_p^2 for use when $N>12$.

One of the simplest arrangements which is usually recognized as "non-random" is one in which each successive term is higher than the preceding. Although such an arrangement would be expected to occur only once in $N!$ arrangements, N would have to be 10 or over to reduce the probability associated with the empirical χ_p^2 to .01. The probability values for $N=6$ and 12 are .4528 and .0924 respectively.

The distributions of turning points and numbers of completed runs are also obtained. It is shown that these distributions approach normality for $N \geqq 12$.

Tests based on these distributions are illustrated by application to annual production, yield per acre, and acreage harvested for sweet potatoes in the

United States from 1868 to 1937. As applied, the test selects only one of these as showing evidence of cyclical behavior even though each one presents evidence of definite primary trends which can be detected by other types of analysis. A basis for choosing one of a group of fitting curves with various values of χ_v^2 is presented.

P. S. Olmstead

Bell Telephone Laboratories

Training and Recruiting of Personnel in the Rural Social Studies, by Theodore W. Schultz assisted by Lawrence W. Witt. Washington: American Council on Education. 1941. xiii, 340 pp. $3.00.

This useful and timely report on the pressing problem of personnel selection in rural social studies was prepared under the direct supervision of the following distinguished committee: Edwin G. Nourse (Brookings Institution), Theodore W. Schultz (Iowa State College of Agriculture), Waldo E. Grimes (Kansas State College of Agriculture), William I. Myers (New York State College of Agriculture), T. Lynn Smith (Louisiana State University), and O. C. Stine (Bureau of Agricultural Economics, U.S.D.A.).

The composition of the committee, representing both Federal and state, both educational and administrative points of view, fitly reflects the need for coordination in studying and in meeting the problems raised by the Report. Rural social studies have experienced a mushroom growth since the last war, and the time for reckoning and for planning is at hand.

While the central problem is that of ensuring a proper quantitative balance between supply and demand in personnel, the committee has found that adequate reform in this line involves reorganization of training programs and promotional practices at the undergraduate and graduate levels and in the official staffs of Land-Grant Institutions and Federal agencies.

At the very basis of the whole program lies the curriculum. Agricultural Economics as a field has never been so well systematized from the standpoint of instruction as some of the older social science disciplines. Undergraduates must gain basic intellectual training in sociological and economic theory, as well as in statistics, history, psychology and other related studies. At the graduate level this discipline must be continued, care being given to correlate field work with theory.

On the staffs of Land-Grant Colleges provision should be made to allow graduates adequate opportunity for leaves of absence. Extension work, experience in both state and Federal agencies, research, intimate knowledge of the problems of two or more different regions should all find their place. The able graduate student must be encouraged to undertake the training which will fit him for future service; he must not be treated as a convenient "assistant" on whom to unload departmental dirty work.

Both in the college departments and in the governmental bureaus efficient personnel practice is vitiated by too much inbreeding. While Civil Service

delays account in part for the hesitancy of government agencies to look "outside" for new recruits, there is also, as in the colleges, a tendency to mistake conformity of method, background and general outlook with prerequisites for efficiency.

The committee concludes its admirable survey with the statement that 50-75 individuals with 3-4 years of graduate training will be needed every year. Measures must be set in motion to regulate the admission of candidates, the elimination of the unfit, and the competency of the survivors.

E. Y. HARTSHORNE

Library of Congress

The Prediction of Personal Adjustment, by Paul Horst and Others. New York: Social Science Research Council, Bulletin No. 48, prepared for the Committee on Social Adjustment under the Direction of the Subcommittee on Prediction of Social Adjustment. 1941. xii, 455 pp. $2.00.

Paul Horst, editor of *Psychometrika*, engaged for some years in personnel selection research for a large industrial concern, and a consistent contributor of statistical articles in the field of mental measurement, was called to Chicago last year to assume charge of the short-time production of a monograph on methods of prediction in the general field of adult adjustment. With the assistance of several workers and the counsel of various committees, the monograph was put together in something over three months. The resulting volume reflects the competence of the contributors, the breadth arising from the mingling of various backgrounds, and the diversity of approach found in different disciplines. No effects of haste are apparent, other than possibly the unreconciled frames of reference of the several writers.

The volume presents both the rationale and technical details of prediction. It is concerned (primarily with respect to illustrative material) with the prediction of success in four areas: vocations, education, marriage, and parole. Part I, comprising two-fifths of the text, presents the logic of prediction. It is relatively simple, straightforward, non-mathematical exposition. It presents, in its eleven chapters, a comprehensive and critical statement of considerations and methods; some of the fundamental material is new, and certainly there is no other single treatment that approaches the compass and definition of this. Part II, comprising some sixty per cent of the text, is technical, and illustrates both the mathematical background, and the application to empirical data, of the methods described in the earlier part. Part II is highly original, ingenious, and resourceful. It marks an advance over methods ordinarily understood and employed.

The reader will find, in addition to the general discussion in Part I, a detailed critical treatment of the case study method as a means of prediction; new light on random sampling in measurement; new methods of solving multiple correlation and regression coefficients; procedures for scaling items, and also attributes; methods for reducing the number of variables; choosing

predictable criteria; simplified ways of handling ratings; the function of "suppression variables," and other topics. Richardson's section on combination of measures is the most realistic treatment of weighting which the reviewer has seen.

The entire volume may be regarded as lying, not solely in the field of prediction, but in the emerging statistical field of combinations of measures to index a complex trait through appropriate composites. The development of this area—accurate, sensitive, and faithful indexing or measurement of variables representing general concepts—is of as much basic importance to those interested in the maturing of statistical science in the social fields as is the somewhat narrower problem of estimating over an interval of time. Though somewhat aside from its avowed purpose, the entire study represents a wholesome and timely recognition of the complexity of traits in social or psychological fields, and the technical developments in the work may be regarded as a contribution to much needed refinement of indexing or measurement in areas that are still somewhat nebulous.

Done in connection with a personnel which varied in interests from case-study methods to almost pure mathematics, the work represents something of an attempt to bring together quantitative and non-quantitative approaches. In this it is only partly successful; the combination is more in the nature of an assemblage than an integration. The transfer, on the part of a writer, from thinking in pure mathematics to thinking in practical fields is not accomplished easily. One feels, without prejudice to the excellence of the mathematical contributions, that certain portions of Part II represent an over-confidence or over-enthusiasm for quantitative methods, even if mechanically performed; the volume is clearly concerned more with the development of theory than with the improvement of immediate practice.

Through its two distinct parts, the volume makes an appeal to both the general and the technical worker. It is made readable by excellent organization, clear exposition, and meticulous editing.

Douglas E. Scates

Duke University

Nation and Family, The Swedish Experiment in Family and Population Policy, by Alva Myrdal. New York: Harper and Brothers Publishers, 1941. xv, 441 pp. $4.00.

Sweden, 1934–39, was the scene of the most dynamic integration of social philosophy, statistical research, and political action yet recorded in human history. The principal components in this synthesis were social-democratic institutions and ideology, academic leadership of high order, population analysis, economic research, biological and psychological studies, and extensive social security, health, housing and welfare programs. The catalytic agent was a new principle, initiated in academic circles, tested in popular discussion, and formulated under governmental auspices. This new principle was a democratic population policy directed toward explicit quantitative

and qualitative objectives through the integration of the interests of people as members of families and as members of the nation. Through reference to this principle and the new focus of national interest in children and the family, the whole structure of Swedish economy was re-interpreted, re-oriented, and to a considerable extent actually re-vamped in legislative and administrative acts. The movement was aimed at long-range objectives, rather than quick and spectacular results. It has been temporarily side-tracked by the necessity of directing all the energies of the Swedish nation to preparations for defense. It can, therefore, be given a pragmatic test only if, and when, an international order is established that will again make possible free experimentation in democratic ways of life.

In *Nation and Family*, Alva Myrdal, who was one of the principal actors in this drama, presents to the English-reading world a broad exposition of its conditions, theory and development. This account of an intellectual revolution in Sweden is likely to have far reaching influence on social theory, the direction of statistical research and political progress in other countries. The general theory, which includes the most comprehensive population policy yet enunciated, is appropriate to any country characterized by advanced technology and democracy. The particular synthesis of various factors in economic and social life effected in Sweden is, of course, not wholly appropriate to any other situation. It will, nevertheless, be extremely interesting to observe the reactions of scientists, administrators, and theorists to the stimulus thus provided for a re-examination of objectives and programs in various fields, e.g. public housing (hitherto characterized by a peculiar dearth of coherent theory), nutrition, employment security, health security, birth control, eugenics, and child welfare. The specific treatment of such topics as economics of homemaking, housing for families, feeding a nation, health for the nation, social security and the social care of the handicapped, opportunities for education, recreation and the family will facilitate its use by specialists in various fields.

The author combines an exposition of theory, frankly related to value judgments which are so far as possible stated explicitly, with an objective analysis of conditions and events. The treatment is well balanced and candid; but it naturally reflects personal interests and preferences. The treatment of general economic theory in relation to population policy has been more fully developed by Gunnar Myrdal in various lectures in this country, including those published in *Population—A Problem for Democracy*. The quantitative materials in *Nation and Family* are presented in brief and popular form—sometimes with a regrettable absence of precise definition. The statistician will regret the lack of an extensive statistical supplement. This need is met, at least in theory, by extensive references to Swedish statistical sources. Moreover, an extensive and precise treatment of some of the materials covered is now available in Thomas's *Social and Economic Aspects of Swedish Population Movements, 1750–1933*. Mrs. Myrdal does present much well selected and significant quantitative material, but her principal contribution

to the statistician is a brilliant conceptual organization of critical problems in their social and political context.

The treatment throws light on many neglected problems in population theory, as, for example, that of the "social substratum." It provides the background for a wide variety of institutional changes. Above all, it is characterized by breadth and intellectual integration. It should prove an antidote to many ill-conceived and short-sighted social reforms. Finally, the book can be recommended to all social scientists for sheer intellectual stimulus.

Frank Lorimer

The American University

The Swedish Collective Bargaining System, by Paul H. Norgren. Cambridge: Harvard University Press. 1941. xv, 339 pp. $3.50.

This is a carefully done description and analysis of the collective bargaining system in Sweden. It sets an excellent example for a pioneer study in this field. There are no really good studies in English of the trade union systems of European countries other than Great Britain. The author has taken one of the most successful of the European systems for probing; he has proved it to be a manageable task for concrete presentation; he has hewed closely to his task and has come out with a scholarly study that is a distinct contribution to the subject of the validity and utility of collective bargaining in modern industrial society.

A somewhat telescoped outline of the volume indicates its coverage: development of the trade organization, of both worker and employer in Sweden; the collective agreements and their contents, more particularly the cost-of-living provisions and the piecework system; the technical making of the agreement and governmental intervention; the subordinate processes of local agreement-making under the general controlling agreement and interpretations of agreement by the Labor Court; and finally, the significance of the system as a whole for the Swedish economy and its probable applicability to the United States. The volume is written with an unusual and refreshing clarity and objectivity, and written only after research and field study in Sweden. It shows a balanced comprehension of social desiderata and economic possibilities.

Professor Slichter's Foreword is in the nature of a review of Dr. Norgren's study. He points out features of interest to both unions and employers in America: the high degree of organization on the part of Swedish employers; the controlling influence of the national trade union federation over its members; the prevalence of national agreements and the absence in those agreements of detailed restrictions limiting output; widespread existence of piecework, which extends into the building trades; safeguarding on both sides of the right to strike, even at the expense of considerable lost time from strikes, which interestingly enough appears to be about as great as in the United States.

Another aspect of the Swedish system is effective intervention of the Government in all industrial disputes and the ready acceptance of it by both employers and workers. Such intervention takes place not only on the level of the interest disputes, but also in justiciable disputes—the author uses the word "interpretational"—which are handled by a special Labor Court. The author might also have mentioned the evolutionary trial and error method by which Swedish disputes legislation has developed. This legislation frequently has been enacted for fixed periods of time, and has had definitely to be reconsidered at the end of such period. The high quality and experience of the mediators have accounted for much of the successful handling of disputes. Nor does the author, it seems to me, emphasize sufficiently the wisdom of the Government and its representatives in weighing "the more objective factors—the swings of the business cycle, the degree of organization on both sides, and so forth." It is this wisdom on the part of the Government that has, I think, made "the effect of mediation on the final outcome of the bargaining process" of more than "minor significance." The Government has not tried to outface economic realities by fiat; neither has it bolstered first one side and then the other; nor has it become an administrative agency of the special interests of either side.

LEIFUR MAGNUSSON

The Library of Congress

Industrial Wage Rates, Labor Costs, and Price Policies, by Douglass V. Brown, Charles A. Myers, John A. Brownell, John T. Dunlop and Edwin M. Martin. Temporary National Economic Committee. Monograph No. 5. Washington. 1940. xxvi, 172 pp.

This study should be examined carefully by anyone engaged in theoretical or empirical work on the economics of the firm. It demonstrates the wealth of economic data available in business records and suggests many ways in which these data may be turned to account.

The first half of the monograph analyzes the experience of two paper companies and two cotton textile companies over the period 1936–38, and of two shoe companies for the years 1931–38. A more intensive case study of the International Harvester Company over the period 1929–37 comprises the second half of the volume. Information is provided on the movement of wage rates, unit labor costs, unit total costs, and prices, on the accounting procedures by which cost figures were derived, and on some of the considerations which appeared to influence price and wage decisions.

The case studies are suggestive rather than productive of definite conclusions. The dominant influence of overhead on the cyclical movement of total unit costs stands out clearly, particularly in the Harvester Company case. The authors suggest that the role of costs in price formation is essentially different in the two major phases of the cycle. "This survey suggests that when sales are falling, business men commonly regard themselves as being forced to lower costs to conform to reduced prices and that, when business is

improving, they think of prices as being raised to conform to higher costs." This lack of symmetry undoubtedly exists, but one may doubt whether the role of costs has been correctly stated. Do business men raise prices in recovery because their own costs have risen? Or are they able to raise prices because the costs of *other producers* have risen sufficiently that all will concur in a price increase?

A number of the firms studied show a systematic deviation of average hourly earnings from wage rates over the course of the cycle, and an increasing secular deviation of unit labor costs from wage rates. The analysis of these deviations in the Harvester Company is particularly interesting. Changes in wage rates and labor costs seem in general to play a very minor role in price formation. Few of the cases studied show any clear connection between the timing and magnitude of wage and price changes.

Research into business decisions must increasingly probe beneath the superficial descriptions which business men give of their behavior and penetrate into the basic operating records of the firm. It is to be hoped that this illustration of what can be done with business records will stimulate additional investigation both by individuals working on one or a few firms and by organized groups able to analyze a considerable sample. One may also hope that this work will be directed increasingly toward answering the questions posed by economic theory in this field, and that we shall in time have a thorough going revision of the economics of the firm.

Lloyd G. Reynolds

The Johns Hopkins University

Hourly Earnings of Employees in Large and Small Enterprises, by Jacob Perlman. Temporary National Economic Committee. Monograph No. 14. Washington, 1940. xv, 94 pp.

Wage data collected in recent years by the Wage and Hour Division of the Bureau of Labor Statistics are here assembled and focussed on the question whether employees of large enterprises enjoy higher average hourly earnings than employees of small enterprises. No clear relation between hourly earnings and size of enterprise was discovered in eight relatively small-scale industries: shoes, leather, radio parts, furniture, hosiery, cotton goods, woolen and worsteds, and knitted underwear and outerwear. Hourly earnings were found to increase with size of enterprise in eight highly-concentrated industries: meat packing, iron and steel, electrical goods, radio sets, explosives, soap, fertilizers, and tobacco products. This tendency existed, moreover, within each region of the country, each size of community, and each grade of labor. A net relationship between size of company and hourly earnings of employees is strongly indicated, though this was not tested by multiple regression analysis.

The qualifications which must be attached to these results are carefully noted by the author. The fact that large and small enterprises nominally in the same industry may not be producing the same product raises difficul-

ties of comparability which can never be entirely surmounted. The distribution of hourly earnings does not give us the distribution of either wage rates or unit labor costs. Where piece rates are used, high earnings simply indicate high man-hour output. Higher earnings in large enterprises might be due very largely to the fact that they make greater use of piece-work and bonus systems which induce a greater input of effort by the worker.

Frequency distributions of hourly earnings for large, intermediate and small establishments in four industries are published in appendixes. The upper limit of the range is usually about the same for small companies as for large, but the lower limit is considerably lower for small concerns. The distributions are markedly symmetrical, but with a tendency toward bimodality in the distributions for intermediate and small producers.

These results suggest many interesting problems. To what extent are the observed differences in hourly earnings attributable to differences in man-hour output rather than in wage rates? Do the higher earnings obtained in large enterprises allow them to skim the cream from the labor market? Are differences in wage rates largely offset by differences in labor efficiency in large and small enterprises? Case studies of large and small enterprises selling in the same product market and buying labor in the same community might furnish very useful evidence on these matters.

LLOYD G. REYNOLDS

The Johns Hopkins University

Economics of Social Security, by Seymour E. Harris. New York: The McGraw-Hill Book Co., Inc. 1941. xxvi, 455 pp. $5.00.

Professor Harris' purpose in writing on this subject is made clearer by the subtitle, which indicates that the volume describes "The Relation of the American [Social Security] Program to Consumption, Savings, Output, and Finance." The book is an outcome of the author's interest in the problem of investing social security funds. He is impressed with the necessity of attaining perspective in evaluating the probable impact of these funds on the whole economy and particularly on the volume and fluctuations of output and employment. As a result, social security, as the author uses the term, means, for the most part, the system of old age and survivors' insurance, with a much lighter emphasis on unemployment compensation and no consideration of other phases of the program.

There are three major divisions, which include a total of twenty-one chapters. Part I considers the relationship of the program to probable output. Part II describes effects of the system on finance. Part III is concerned with the incidence of costs of the program.

In each major division, there is an extensive citation of varying viewpoints, well organized, logically arranged, clearly quoted or paraphrased. A wide range of pertinent statistical data is included, and footnote references provide a fairly comprehensive bibliography of recent literature on economic implications of social security measures. Major divisions and chapters are

effectively arranged, in that varying viewpoints are first contrasted, after which the author seeks to strike a balance in a concluding section or paragraph. As a result, the presentation sustains the reader's interest and stimulates a critical viewpoint, which, in the opinion of this reviewer, gives the volume its greatest value and justifies the opinion that it is a "must" on the reading list of all those who are charged with responsibility for administering these features of the social security program. Because of this arrangement, also, the book will be useful in the classroom, where opposing viewpoints may be criticized in more detail.

In other words, Professor Harris has provided a powerful incentive for thinking critically of the probable effects of our old age security program. He repeatedly notes, for instance, the possibility if not likelihood of the repudiation of obligations under old age provisions. Indeed, the volume is replete with suggestions, implications, and possibilities, and they represent its most important contribution. For the conclusions to which the author comes are, as would be anticipated in a study in which so much emphasis is placed on authority and so little on pertinent evidence, distinctly controversial. That is particularly true, also, because the author attempts to juggle and balance a coterie of economic tendencies, actuarial statistics, and principles of social justice. In effect, the reasoning process attempts to follow the threads of an endless number of variables through a complex maze of arguments and thereby to arrive at a series of tenable conclusions. By such a process, it is possible to formulate hypotheses, but anything in the nature of a positive conclusion requires more tangible evidence. Illustrations of this point are numerous. After one hundred fifty pages of citation and quotation, the conclusion as to the incidence of costs is far from decisive. The conclusion as to the impact of the program on prices is a series of questions. The author's decision that, on balance, "the stagnationists have a strong case" is of similar caliber, as is the declaration, in curious scientific language, that "our sympathies are with the spenders."

Aside from its value in focussing attention on the place and importance of social security in the total economy, therefore, it appears that the major value of the book is its profusion of illustrations of the need for empirical studies to supplement armchair hypotheses.

DALE YODER

University of Minnesota

Federal Finances in the Coming Decade: Some Cumulative Possibilities, 1941–51, by Carl Shoup. New York: Columbia University Press. 1941. vii, 121 pp. $1.00.

Essential Facts for Fiscal Policy. New York: National Industrial Conference Board. 1941. 135 pp.

It has often been said that although the Treasury plans only from year to year, it is long range planning of finances that is required. Professor Carl Shoup has boldly and successfully attacked the problem of long range fiscal

planning. That his estimates, or better, his informed guesses may prove to be wrong, is of secondary importance. In fact Pearl Harbor has already dated many of his findings. What is of crucial importance is his method of approach, his assumptions, his selection of significant variables and the expert use of series which throw light on this problem. Economists, civil servants and statisticians will, I am sure, be grateful to Professor Shoup for this excellent monograph.

For the next 10 years he estimates expenditures, tax receipts inclusive of yields associated with five successive revisions in the tax structure, each in turn contributing additional amounts of taxation, and the rise of debt in the absence of tax revisions and on the assumption of these five successive changes in the tax structure. The author does not neglect the relation of taxes to income, of taxes to prices, of the rise of yield of one tax upon the yields of others, of the effect of high tax yields on sales of securities.

Professor Shoup is concerned almost exclusively with the fiscal aspects of the problem; but he is nevertheless anxious that fiscal policy be used as a means of keeping prices from rising. It is possible, of course, to deal with the price problem exclusively through controls. The cost then would be a complicated and comprehensive administration, regimentation, and a rapidly growing public debt. The price problem would be solved at the expense of the public debt. It would have been helpful if Professor Shoup had suggested the relative contributions to be expected of taxes, controls, and monetary policy. Washington needs guidance here.

In discussing the sale of government securities, Professor Shoup raises many interesting questions. What is the magnitude of probable sales to institutions and individuals? What limitations are placed on sales by banking capital and cash reserves? What is the elasticity of supplies of loanable funds in response to a rise in the rate of interest? In particular, will the cost to the government of the rise in the rate of interest exceed the new money attracted by the increase of the rate on government bonds? On these and many other issues Professor Shoup sheds new light.

He is perhaps overanxious concerning the limits on sales of bonds to banks set by cash reserves and capital. The monetary authority can deal adequately with these problems; and the danger of rising rates of interest on the value of assets and hence on the adequacy of the capital structure can be averted through Treasury guarantees against depreciation of securities. Advances in this field are beginning to be made. Professor Shoup might also be encouraged by British experience in the last two years: Vast drains on the capital market have been accompanied by a decline in the rate of interest on public securities. In almost two years of warfare the price of all fixed interest yielding securities rose by almost 15 per cent. The demand for government bonds, observe, stems not only from existing supplies of money, but also from new supplies; not only from normal savings, but also from additional (or diverted) savings induced by higher incomes and limitations on spending both for consumption and capital purposes. Through a control of

the supply of money and the demand for cash for spending, the government can assure the Treasury a reasonable rate of interest. And banks should not be guaranteed a rate in excess of the variable costs involved in absorbing additional public debt; and when outlets for cash increasingly are closed private purchasers will be pleased with the opportunity to purchase government securities at 2 per cent.

Professor Shoup's alternatives for increasing public revenues are on the whole eminently sensible. He suggests under his first two alternatives large rises in the income taxes and other taxes on capitalist income; under the third, excise duties; under the fourth, manufacturers' sales taxes; and under the fifth, payroll taxes. The respective additions to receipts for the fiscal year 1944, in which year the total gains would be a maximum, are under these successive alternatives $3.7, $1.3, $1.0, $1.3 and $0.8 billion. Total revenue would then rise to $21 billion in the fiscal year 1944. At an income of $100 billion the total tax load for all governments would then be roughly 30 per cent of national income. In the absence of large curtailments imposed on industry by priorities, rubber shortages, etc., I would have suggested that Professor Shoup underestimates the tax potential in the excise field. The reviewer would also suggest a greater increase in payroll taxes than that proposed by Professor Shoup. Despite the increased burden of income taxes, a rise of payroll taxes from the present low level by at least 4 per cent seems eminently desirable. The rise of yield would then be of the order of $2 billion annually. (I assume broader coverage than is now had under the Social Security Act.)

The National Industrial Conference Board Study deals with somewhat similar problems. They apparently would rely heavily on a rise of income tax induced through a lowering of exemptions. Estimates of yield are based on the assumption that the yield of surtax will be given by the existing ratio or normal tax yield to total income tax yield. This assumption has, of course, been violated by the 1941 Revision which relied heavily on a rise of the surtax yield. The Conference Board study also underestimates the potential yield of corporation taxes though a final answer depends on the results of controls on corporate income. Its estimate is that the maximum available draft on corporate income (prior to the 1941 Act) was $1.6 billion. The 1941 act has already increased the tax burden on corporations by $1.4 billion. Professor Hart estimates the corporate tax base in the period just ahead at $9 billion (exclusive of inter-corporate dividends). This figure may be low, but in any case, the tax assessment even now is but $1 billion. At higher income levels in 1942, the rise of corporation taxation may well yield several billion additional. In 1937, the distributable net income was $10.2 billion, cash dividends $7.3 billion and taxes $1.4 billion.

Again, a suggested limit of $\frac{2}{3}$ to $\frac{3}{4}$ for the excess profits tax seems too low. The American experience in the last war and current British experience with 100 per cent tax (and possible rebates of 20 per cent in the post-war period) give an indication of the possibilities. Business profits are today

largely determined by government demand. A satisfactory market is assured for the vast majority of market procurements. A seller's market of this type calls for heavy excess profits. Adverse effects on motivation are not likely to be serious. The Conference Board study also underestimates the possibilities of taxes on luxuries. Our experience in the last war and the twenties is not conclusive on this score. Unsatisfactory yield will be explained by a factor not given adequate attention by the Conference report: curtailment of production. But the government has not adequately explored the possibility of supplementing scarcity with high prices inflated by severe taxes. This is the correct procedure for commodities on the luxury list.

Not satisfied with an elaboration of the fiscal aspects of the problem, the Conference Board also estimates the amount of real resources available for the defense effort. Unfortunately the real and financial aspects are not integrated well. Furthermore, the estimates of optional consumption and expenditures above practical minimum level raise serious questions of definitions. "As in the case of minimum expenditures, efficiency expenditures are increased as income increases on the basis of food expenditures"[1] (p. 25). Optional expenditures are the excess of total expenditures over efficiency expenditures, which are, as is indicated above, obtained in a very crude manner. Again, we are informed that optional consumption in 1937 for automobiles, for example, was 25.9 per cent of total expenditures. Obviously the optional ratio will be larger for short periods than for long periods. Expenditures for automobiles may well be cut by 75 per cent in the next few years. Another troublesome point is the interrelation for groups of items of the percentages of optional to total expenditures. Will consumption of each class of items be cut equally if the reductions are applied to all classes as if they were applied to but one or a few classes? Does a reduction of 20 per cent in automobile expenditures make more difficult or less difficult a reduction of x percentage in recreation?

In this brief review, the writer has noted several points of disagreement with the Conference report. He is not, however, unaware of the many points of agreement, the rich collections of materials made available by the resources of the Board, the excellent summaries of tax and debt history, and the interesting study of income, consumption, and sales. No student of our present fiscal problems can afford to do without this volume.

SEYMOUR E. HARRIS

Harvard University

Air Transportation, by Claude E. Puffer. Philadelphia: The Blakiston Company. 1941. xxiv, 675 pp. $3.75.

This book is primarily an economic description and history of the air transportation industry in the United States. It is well written and very

[1] "The ratio of total expenditures ($1316) to food expenditures at the efficiency level ($487) was 2.7. The efficiency level would then be $617×2.7 or $1666." (Food expenditures at this income are estimated at $617.)

comprehensive, and hence fulfills an urgent need of present students of the subject. The author applies the typical demand and cost analyses of economic theory to the services offered by the air carriers—passenger, mail, and express. The study includes a well-organized discussion of the legal characteristics of the industry, the functions of the Civil Aeronautics Board, intercorporate relationships, labor relations, and safety regulations.

Only one major criticism appears to the reviewer. The author chose to follow the lead of Paul T. David and Joseph B. Eastman in his analysis of the public aid included in air mail payments. This school of thought contends that such aid should be measured by the excess of Post Office disbursements to the airlines above the airlines' cost of rendering the required service. The trouble with this line of reasoning is that it leads to some peculiar conclusions. It could be concluded, for example, that one way to eliminate such public aid would be to increase the costs of the airlines to the extent that there would be no excess of disbursements over costs. It could also be concluded that the airlines' cost of rendering the required service could be increased by any amount named and there would be no increase in public aid if the Post Office disbursements were increased by a like amount. The reviewer holds that such aid is the excess of Post Office disbursements over net revenue from air mail.

The magnitude of public aid contained in the air mail payments of the past is really of little significance in the future of this industry, but the author used this same type of analysis in his discussion of rate-making, which is a question of primary importance.

He states that the basic consideration in determining mail payments "should be the cost [of performing mail service] incurred by the carriers (including a reasonable profit), irrespective of what the government receives in postal revenues." At this point the author parts company with Mr. Eastman who said of one of his own works, "this study is not intended to yield the type of cost analysis required for rate-making purposes." In the opinion of the reviewer, a sound rate-making policy for the airlines cannot ignore the air mail revenue of the Government.

Harold J. King

University of Pittsburgh

Air Mail Payment and the Government, by F. A. Spencer. Washington: The Brookings Institution. 1941. xii, 402 pp. $3.00.

This book is a clean-cut job. The author runs through the periods of air mail transportation without getting lost in detail, and yet with convincing thoroughness. An astonishing amount of information is presented about the air mail; in fact, about the air transportation business. There is a good balance between the review of policies under the several governmental bodies in control at one time or another and the criticism of the policies. Where the lapse of time permits, historical judgment of adequacy is used. This is aptly applied to the period of Interstate Commerce Commission rate making.

Controversies are analyzed with apparent fairness. The Post Office Department and the Interstate Commerce Commission difficulties in joint administration are well stated. Likewise, the slow progress of the Civil Aeronautics Authority and the Civil Aeronautics Board in determining sound over-all policies for air mail pay is criticized. At the same time, the author analyzes the problem and finds its complexities great. There is the "subsidy" question, sufficient in its implications for an accounting convention of outstanding interest. Differences between the criteria for public and private expenditures could be more fully analyzed than by the author. These have recently been well discussed by the National Resources Planning Board in its report on *The Economic Effects of the Federal Public Works Expenditures, 1933-38.*

Efforts to rationalize governmental aid by the profit criterion of business seem unlikely to lead to constructive policies. While the major criticisms of the author are appropriately directed at what seem to be governmental faults or deficiencies, he does not become counsel for the air mail contractors. His examples of watered stock, bad financing, and high costs are clearcut and specific. Specious arguments by operating companies for charging the mail with excessive fractions of gross operating charges are refuted.

Perhaps the most informative section for the reader interested in current practice is the extended chapter "Air Mail Compensation under the Civil Aeronautics Act." The rate policy section of the Act and its possible interpretations are well stated. Not merely "need" for compensation must be considered, but "a policy which encourages managerial efficiency."

Finally, the rapid growth of patronage of airlines leads the author to conclude, "If the present upward trend . . . continues, the Civil Aeronautics Board will soon be faced with the problem of deciding whether rates of air mail compensation or passenger and express rates shall be reduced (and, if so, in what proportion) in order to check unreasonable profits accruing to the carriers."

BALDWIN M. WOODS

University of California

Statistical Activities of the American Nations, 1940. Edited under the Direction of the Temporary Organizing Committee of the Inter-American Statistical Institute, by Elizabeth Phelps. Washington: Inter-American Statistical Institute. 1941. xxxi, 842 pp. $2.00.

Not since Koren's well-remembered *History of Statistics,* which commemorated the seventy-fifth anniversary of the American Statistical Association, has there appeared a more useful *aperçu* of the statistical set-up in a number of countries than the present volume, described in the sub-title as "a compendium of the statistical services and activities in twenty-two nations of the Western Hemisphere, together with information containing statistical personnel in these nations." Originating like Koren's in a commemoration—this time the twenty-fifth session of the International Institute

of Statistics scheduled to be held in Washington, 1939 (deferred to 1940 to coincide with the Eighth American Scientific Congress), with which a partial celebration of the American Statistical Centenary was to have been conjoined—the outbreak of war transformed both occasion and auspices. In the event, the International Institute and American Statistical sessions were cancelled; in their stead a statistical section was added to the Scientific Congress for a purely inter-American program of statistical papers, with the idea also of exploring the possibilities of a permanent inter-American statistical organization of professional character. The latter has since taken form as the "Inter-American Statistical Institute." The original "arrangements committee" likewise continued its work on the compendium which is now completed. Thus the purpose of the volume, which was to promote a north-and-south acquaintanceship in the world of statistics as supplementary to the predominant east-and-west flow of the past, has not only been retained, but in effect sets seal upon an epochal movement in international organization and relationships.

The treatment of subject matter is purely descriptive. For each of the twenty-two countries (alphabetically arranged) an authentic article by a local author was secured (for the United States, 23 such "author's articles" were necessary). This is printed in the language of the country (English, Spanish or Portuguese). Preceding each is a "summary" in all three languages. The plan of both article and summary calls for a review of (1) educational facilities, (2) library facilities, (3) statistical societies or associations, (4) non-official or semi-official statistical agencies, and (5) the official statistical system, its organization, progress to date, publications, etc. Though there is a natural tendency in such writing to "put the best foot forward," and though the efficiency of a statistical system must be measured only by its statistical results, the usefulness of such a survey for purposes of reference, especially to official statisticians, need not be stressed.

A striking revelation is the extent to which the Latin American nations have enacted legislation especially in recent years looking to the central control of statistics. At least a dozen, including all the leading ones, have Central Statistical Bureaus. Departmental compilation and publication, however, remain paramount in most. Statistical coordination in fact demands an amount of spade work in advance that is apt to be underestimated. But the beginning lies in the legal frame-work, and in this the South Americans almost without exception are well equipped.

The Directory of Statistical Personnel which concludes the volume lists the leading statisticians of each country (other than the United States) after the manner of *Who's Who*, an admirable introduction to the broader acquaintanceship which is to be.

For the improvement of economic relations throughout the American continents a compendium of this kind was a prerequisite. When prior to World War I a Royal Commission was appointed to explore similar possibilities within the British Commonwealth of Nations its very first recom-

mendation looked to a Conference of Statisticians to review the incompleteness and inconsistencies in official statistics which lay as a stumbling block across the threshold. Conferences in due course took place in London in 1920 and in Ottawa in 1935 with results of first-class benefit. Implicit in the present volume is the development of the Inter-American Statistical Institute as a necessary preliminary and impetus to a similar rapprochement in this Hemisphere.

ROBERT H. COATS

Dominion Bureau of Statistics

Industrial Corporation Reports (76 individual reports for leading manufacturing industries, covering operations during 1938–1939.) Federal Trade Commission, Washington, D. C. 1941.

This set of 76 reports on selected manufacturing industries resulted from the Federal Trade Commission's project for the collection of annual financial data from a large number of corporations operating in some principal manufacturing industries in the United States.

In each report there are presented financial data for 1938 and 1939, comprising a balance sheet, income and expense statement, and an earned surplus statement; also ratios and percentages derived from the basic data. The principal facts in the exhibits of data are domestic and foreign sales; costs and expenses in detail; investment, profits, and rates of return; dividends paid and net income retained by the corporations; operating ratios; inventories; important financial ratios; and increases and decreases in principal assets and liabilities. The 1939 information was obtained from financial reports submitted to the Commission. The 1938 information was procured from annual reports of public record made by the corporations, or from balance sheets submitted to the Commission; the 1938 data are not complete in every instance. All data shown for the several corporations canvassed in each industry are combined in a manner which does not identify the reports from any individual corporation.

The designation of each industry canvassed was made to conform with the "Standard Industrial Classification" used by various departments of the Federal Government. This makes possible ready reference to other statistics released by government agencies, particularly those of the Bureau of the Census.

The reports are limited to a presentation of basic data in terms of the amount of money value with only simple analysis. The scope of the work did not permit the Federal Trade Commission to make detailed adjustments of the reports of the various corporations to make accounting treatment and procedure entirely comparable. The possible variations, however, are held to be unimportant. All data follow the classification of accounts as generally reported by the corporations. Reclassifications in some instances were made by the Commission for the purpose of applying uniform accounting princi-

ples, or where other differences in the accounting method would have affected the comparability of the combined statistics.

There are some discrepancies in comparability with Census data that are explained. The total amount of sales for the limited number of corporations which were canvassed in nine of the 76 selected manufacturing industries exceeded the total value of products for each respective industry as reported by the Bureau of the Census. The reason for this is that Federal Trade Commission data pertain to complete operations of corporations while Census data are by establishments, and occasionally the operations of some plants of a given corporation have undoubtedly been reported in different industry classifications. Such discrepancies occur whenever the corporation operates a number of plants manufacturing different products, and also whenever it operates plants in foreign areas. In the latter case, the Bureau of the Census does not receive reports from foreign establishments of domestic corporations. The same discrepancy probably appears to a lesser extent in many more of the industry reports. Some quantitative indication, or correction if possible, of the amount of foreign operations and of included operations pertaining to other industries than the one reported should improve the future industrial reports.

As far as can be learned from using the statistics, the reports seem to be very carefully prepared, for practically no inconsistencies in data were found between the various parts of the publications and their exhibits. One exception was found in the summary of exhibits for the food specialty manufacturing industry, where the total costs and expenses figure apparently did not include selling and advertising, administrative and general office, taxes, and research and development expenses.

The reports fill a need in supplying statistical data on the financial operations of important industries and should be valuable to economists, statisticians, businessmen, and others interested in the productive activity of industrial corporations and in their financial results. The Commission's claim that much of the information in the reports is not elsewhere available in the same detail is undoubtedly true, and such reliable information about operating conditions may serve as a guide for business management. The Federal Trade Commission plans to publish similar industrial reports annually.

OSCAR L. LERNER

National Resources Planning Board

PUBLICATIONS RECEIVED*

American Statistical Association, Albany Chapter, Central Statistical Committee. *Index of Statistical Data Available in New York State Departments.* New York: State Division of Commerce, Bureau of Planning, Executive Department. 1941. xx, 107 pp.

Brumbaugh, Martin A., and Kellogg, Lester S., with the collaboration of Graham, Irene J. *Business Statistics.* Chicago: Richard D. Irwin, Inc. 1941. xv, 913 pp. $4.00.

Burnham, James. *The Managerial Revolution.* New York: The John Day Company, Inc. 1941. 285 pp. $2.50.

Clarkson, Jesse D., and Cochran, Thomas C. *War as a Social Institution.* New York: Columbia University Press. 1941. xvii, 333 pp. $3.50.

Coit, Charles Gurtney. *The Federal Reserve Bank of Richmond.* New York: Columbia University Press. 1941. xv, 140 pp. $2.00.

Dean, Joel. *The Relation of Cost to Output for a Leather Belt Shop,* with a Memorandum on Empirical Cost Studies by C. Reinold Noyes. Technical Paper 2. New York: National Bureau of Economic Research. 1941. 74 pp. 50 cents.

Doherty, Richard P. *Trends in Retail Trade and Consumer Buying Habits in the Metropolitan Boston Retail Area.* Boston: Bureau of Business Research, Boston University, College of Business Administration. September, 1941. 40 pp. $1.00.

Ellis, R. K., Davies, George R., and Yoder, Dale. *Workbook for Business Statistics.* New York: John Wiley & Sons, Inc. 1941. 40 pp. $1.00.

Ezekiel, Mordecai. *Methods of Correlation Analysis* (Second Edition). New York: John Wiley & Sons, Inc. 1941. xix, 531 pp. $5.00.

Fisher, R. A. *Statistical Methods for Research Workers* (Eighth Edition). Edinburgh: Oliver & Boyd. 1941. xv, 344 pp. 16 shillings.

Glover, John George, and Madden, John T. *The Development of American Industries.* New York: Prentice-Hall, Inc. 1941. xxviii, 1005 pp. $5.00 (School, $4.50).

Holzinger, Karl J. and Harman, Harry H. *Factor Analysis.* Chicago: University of Chicago Press. 1941. xii, 417 pp. $5.00.

Horst, Paul, et al. *The Prediction of Personal Adjustment.* New York: Social Science Research Council. Bulletin 48. 1941. xii, 455 pp. $2.00.

Jacobstein, Meyer, and Moulton, Harold G. *Effects of the Defense Program on Prices, Wages, and Profits.* Washington: The Brookings Institution. 1941. 43 pp. 25 cents.

Kuznets, Simon, assisted by Lillian Epstein and Elizabeth Jenks. *National Income and Its Composition, 1919–1938.* New York: National Bureau of Economic Research. 1941. Two volumes. xxx, 950 pp. $5.00.

League of Nations. *World Economic Survey, Ninth Year 1939/41.* Geneva Princeton: Princeton University Press. Economic Intelligence Service. 1941. 275 pp. $2.50.

Lundberg, George A. *Social Research* (Second Edition). New York: Longmans, Green and Co. 1942. xx, 426 pp. $3.25.

Minnesota, University of. *Proceedings of the 23rd Annual Meeting of the American Association of Collegiate Schools of Business.* Minneapolis:

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. Some items of minor interest to statisticians have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications which are to be found in the *Revue de l'Institut International de Statistique, Journal of the Royal Statistical Society, American Economic Review, Population Index, Transactions of the Actuarial Society of America, The Record of the American Institute of Actuaries,* and *Sankhyā, The Indian Journal of Statistics.*—Editor.

School of Business Administration, *University of Minnesota.* 1941. 135 pp. $1.25.

Myrdal, Alva. *Nation and Family.* New York: Harper & Brothers Publishers. 1941. xv, 441 pp. $4.00.

Nicholls, William H. *Imperfect Competition within Agricultural Industries.* Ames: The Iowa State College Press. 1941. xiv, 384 pp. $3.75.

Pancoast, Omar, Jr. *Occupational Mobility.* New York: Columbia University Press. 1941. ix, 155 pp. $1.75.

Rich, Bennett Milton. *The Presidents and Civil Disorder.* Washington: The Brookings Institution. 1941. x, 235 pp. $2.00.

Ross, Murray. *Stars and Strikes, Unionization of Hollywood.* New York: Columbia University Press. 1941. x, 233 pp. $2.75.

Shoup, Carl. *Federal Finances in the Coming Decade, Some Cumulative Possibilities, 1941–51.* New York: Columbia University Press. 1941. xii, 121 pp. $1.00.

———. *The Prospects for a Study of the Economic Effects of Payroll Taxes.* Washington: Committee on Social Security, Social Science Research Council. 1941. v, 74 pp.

Special Libraries Association. *Special Library Resources*, Edited by Rose L. Vormelker. New York: Special Libraries Association. 1941. $6.00.

Spencer, Francis A. *Air Mail Payment and The Government.* Washington: The Brookings Institution. 1941. xii, 402 pp. $3.00.

Tax Foundation. *Tax Facts & Figures.* New York: The Tax Foundation. 1941. 73 pp.

Tax Institute. *Tax Yields: 1940.* Philadelphia: Tax Institute of the University of Pennsylvania. 1941. x, 149 pp. $2.50.

Wallis, W. Allen, and Moore, Geoffrey H. *A Significance Test for Time Series and Other Ordered Observations.* New York: National Bureau of Economic Research, Technical Paper 1. September 1941. xii, 59 pp. 50 cents.

Waterman, Merwin H. *Economic Implications of Public Utility Holding Company Operations.* Ann Arbor: Michigan Business Studies, Vol. IX, No. 5. Bureau of Business Research, School of Business Administration of the University of Michigan. 1941. 175 pp. $1.00.

Whitsett, J. M. *Banking Operations in Ohio, 1920–1940.* Columbus: The Ohio State University. 1941. xxv, 217 pp.

Wickizer, V. D., and Bennett, M. K. *The Rice Economy of Monsoon Asia.* Stanford University: Food Research Institute. 1941. xiii, 358 pp.

Woodbury, Robert Morse. *International Comparisons of Food Costs.* Montreal: The International Labour Office, McGill University. 1941. ii, 78 pp. $1.00.

Government Publications

Department of Commerce, Bureau of Foreign and Domestic Commerce. *The Balance of International Payments of the United States in 1940*, by Hal B. Lary and Paul D. Dickens. 1941. vi, 93 pp. 20 cents.

Executive Office of the President, Bureau of the Budget. *Standard Industrial Classification Manual.* Volume I. Manufacturing Industries. 1941. xiii, 00 pp. 15 cents.

Federal Deposit Insurance Corporation. *Municipals.* Washington: Committee on Municipal Obligations, National Association of Supervisors of State Banks. 1941. 108 pp. $1.00.

National Resources Planning Board. *Family Expenditures in the United States.* June, 1941. xxii, 209 pp. 50 cents.

Work Projects Administration, Division of Research. *Mexican Migratory Workers of South Texas*, by Selden C. Menefee. 1941. xvi, 67 pp.

State of New York, Division of State Planning. *Recent Trends in Important Manufacturing Industries in New York.* Albany. 1941. iv, 341 pp.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 218 JUNE, 1942 *Volume 37*

ON A CLASSIFICATION OF THE PROBLEMS OF STATISTICAL INFERENCE

By W. Edwards Deming
Bureau of the Census

This paper springs from an attempt to set forth the factors that give rise to scientific decisions for action, and to show how the results of experiments and surveys (in other words, the numerical data of nature) are used as part, but only part, of the evidence required for these decisions. An analysis of the factors that influence scientific decisions and recommendations for action, and an attempt to state the responsibilities of the statistician, bring about a useful classification of the problems of statistical inference. This classification has been helpful in the design of enumeration and tabulation procedures, both for samples and complete counts, and it applies equally well in industry and in the natural and social sciences. The usefulness of mathematical statistics, or of any other method of inference, will be measured by its ability to assist judgment in making better predictions, and making them oftener. The classification made here *provides a guide for evaluating the usefulness of mathematical statistics in specific problems, on the grounds just* stated.

DATA ARE FOR PREDICTION AND ACTION

The ultimate purpose of taking data is action. Scientific data are not taken for museum purposes; they are taken as a basis for doing something. If nothing is to be done with the data, then there is no use collecting any. The ultimate purpose of taking data is to provide a basis for action or a recommendation for action. The step intermediate between the collection of data and the action is prediction.

Every empirical statement of science is a prediction. It is a philosophic commonplace that every empirical statement in science has temporal spread, and partakes the nature of a prediction. There is no scientific interest in any measurement or empirical relationship that does not help to explain what will happen or has happened at another time or another place.[1]

[1] C. I. Lewis, *Mind and the World-Order* (Scribners, 1929), pp. 129, 132, 195. A mathematical result is not put to a test by what actually happens in nature; it is not empirical.

An operation of measurement (experiment or survey, carried out by sampling or complete count), if it is repeated over and over, creates a sequence of results, commonly known as a *population* of measurements. Any one measurement is but *one term in a sequence of terms*, and this sequence of terms actually or theoretically might be extended by repeated applications of the operation. Not this one term, but its relation to the *rest of the sequence* is the point of interest. When you say that the length of this table is 6 feet you make a prediction: you imply that anyone repeating this or any accepted method of measurement will also find the length of this table to be 6 feet within limits that you must specify, depending on the requirements.

The publication of a measurement is in two ways a prediction with regard to measurements *not yet taken*; first, it is a prediction with regard to repeated applications of this one method (more terms of *this* series); and second, by implication at least, it is a prediction with regard to repeated applications of other methods (terms of *other* series).

The announcement of a functional relationship is likewise a prediction. A curve fitted to a set of points is of interest, not on account of the data fitted, but because of data *not yet fitted*. How will this curve fit the *next* batch of data?

ACTION, EVIDENCE, AND THE STATISTICIAN'S JOB

What constitutes evidence? What constitutes a statistical method? When a problem arises, demanding action, action will be taken. The scientific attitude is to base the action on rational predictions and the degree of belief associated with these predictions, as well as on the possible consequences of different courses of action. The degree of belief in any prediction will depend on the evidence available: a change in evidence will change the degree of belief, and hence ultimately, possibly the action also. The amount of money that should be spent collecting evidence depends on the hazards of the action. In some circumstances, indications afforded by a scant amount of data will suffice as evidence for the action required. In more hazardous circumstances, a vast amount of data may be needed as a basis for action. Of course, there is not always time to collect the evidence that is really needed, as when action must and will be taken at a certain time, whether or no.

Information that does not affect the degree of belief is irrelevant to the purpose and is not evidence. Likewise, any method of inference that does not help to predict, or which does not affect the degree of belief in some prediction that needs to be considered, is irrelevant to the purpose. A method that is useful in one set of requirements may not be in another.

An experiment or survey (such as a census) should be designed to provide evidence on which to evaluate the degree of belief in the predictions that are assumed to be useful in formulating a course of action for a problem that is faced. A single experiment or survey, however, will rarely if ever constitute all of the available evidence. No evidence from other experience dare be ignored if it affects the degree of belief in a prediction that needs to be considered in deciding an important course of action.

Action tests the prediction. If there is to be no action, and hence no test, then any prediction, regardless of evidence, may safely be made. On the other hand, if the consequences of taking the wrong action will be costly in time, money, materials, comfort, or prestige, then the consequences of the action will be carefully weighed along with the degree of belief and the evidence for the prediction. The dependence of action on prediction, the degree of belief associated with the prediction, and the consequences of the action, can be put in the accompanying diagram.

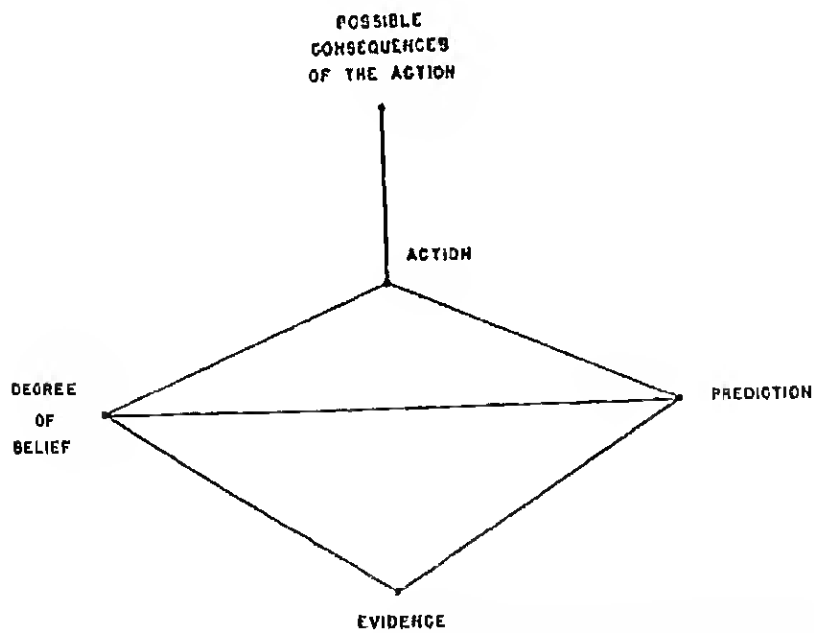


Diagram showing that the action adopted depends on the possible consequences of the action, and on a prediction and the degree of belief associated with the prediction. The degree of belief in a prediction depends on evidence. The lower triangle is the Shewhart diagram of the three components of knowledge—evidence, prediction, and degree of belief.

The statistician's job. It is the statistician's job to make rational predictions concerning measurements yet to be made. He also helps other people make predictions by providing evidence (e.g., collecting data), and evaluating this evidence. Anyone who makes a statement that is based partly or wholly on data of nature is in some respect a statistician, but the statistician is expected to make better predictions oftener than other people and to know what confidence he may have in his predictions; that is his business. Since action will be based on his predictions, and thus put them to a test, he is vitally interested in the consequences of his predictions, hence in the evidence and degree of belief in them. He cannot afford to be wrong in the wrong place. A prediction that cannot be put to a test, at least theoretically, or which is not intended to be put to a test, is not a scientific statement, and does not require a statistician; anyone can make such predictions. The statistician's job therefore appears to be fourfold:

i. To plan the collection of data to provide evidence for whatever predictions must be considered in the decision for action.
ii. To describe the method by which the data were collected, and to present the data by summaries and comparisons. This presentation must be carried out so that the degree of belief, and hence the action taken, will be the same on the basis of the summary as it would be on the basis of the original data.[2]
iii. To make predictions as a basis for action. This will require knowledge of the subject matter, and perhaps mathematical statistics (vide infra).
iv. To make recommendations for action. In so doing, the statistician will take account of all available evidence and the degree of belief in the predictions that need to be considered; he will take account also of the consequences of each possible course of action.

TWO TYPES OF STATISTICAL PROBLEMS

Description of the two types. In my own work I have found it useful to distinguish between two types of problems that confront the statistician in his job of making predictions:

Type A. Problems in which action is based on a prediction regarding future measurements of a product *already in existence*. The evidence comes from measurements already made on the thing itself, or on samples thereof.

[2] This is merely a restatement of Shewhart's rule 2, given in his *Statistical Method from the Viewpoint of Quality Control* (The Graduate School, Department of Agriculture, Washington, 1939), p. 92.

Type B. Problems in which action is based on a prediction regarding future measurements of a product *not yet subject to measurement*, perhaps not even produced yet. The evidence comes from measurement made on other product originating from the same or similar processes.

In both types of problems the ultimate purpose is action. In problems of Type A, some action is to be based on the measurements of a product, or samples thereof. The Type A problem might be described as the problem of measuring something. *Interest centers in the product as it is*, not on how it got that way, or what it might to be or might have been. The action taken may be, and often is, action on the thing that is measured (see the example described below). The evidence for the prediction is furnished by (i) measurements that have been made on this particular product; (ii) previous experience with the method of measurement that is used; and (iii) previous experience with the method of sampling, if sampling is used.

In contrast, in problems of Type B, the action is based on predictions regarding future measurements of some product not yet subject to measurement, perhaps not even produced yet. *Interest centers in the process*, the underlying cause system of forces (social, economic, mechanical, physiological, chemical, geological, biological), that give rise to yesterday's, today's, and tomorrow's product. It is through a study of the underlying process and what it has produced in the past that one is able to predict the product of the future - more generally, the results of measuring a product not yet subject to measurement. Studies of Type B are carried out with the aid of measurements made on past product, originating from basically the same or related processes. In some lines of work (notably industry and agriculture) these studies may lead to a modification of the process itself.

An example. The distinction between Type A and Type B can be illustrated by consideration of an industrial product. Each lot is inspected in order to determine what disposition shall be made of it. Disposition of the product will be made in accordance with certain rules of action, action on each lot being taken, we shall presume, on the basis of the quality of that lot and that lot alone, regardless of what its predecessors have been. The disposition may be to pass it, accept, reject, regrade, sort, or repair it. This is a Type A problem, because the action depends on what the lot is, not on what it might have been, or what is expected of future lots. The easiest and cheapest way to inspect the lot may be to measure all of it: if this is too costly or otherwise im-

practicable, sampling will be used, possibly with elaborate sampling design, calculation, and experimentation. Schemes of double sampling may be helpful.[3]

The lots are inspected for another purpose also. It is desirable to keep an eye on the production process, to forestall an inordinate percentage of rejections *in the future*, and to point the way to effecting improvements or desired changes in future quality. This is a Type B problem, and would be carried out by every possible means to attain the end. Investigations will be made into the underlying mechanical and chemical forces that make the product what it is. The quality of past product, made available by Type A investigations, will be brought into the study. It may be profitable to watch the hourly variations of quality, as observed in samples of current product (cf. the section "Some remarks on quality control in industry"). Such studies are directed primarily for *action on the process*, not for the disposition of any particular lot of product.

Similar analogies can be drawn from social and economic studies. Type B problems constitute a large part of science, natural and social. The establishment of any causal or functional relationship, for instance, is a Type B problem. The ultimate goal in establishing a relationship is to arrive at a theory or formula that will hold, with stated limitations, for data not yet taken. Varietal and treatment tests in agriculture, for instance, are not made just to determine which *was* best under certain conditions (Type A), but rather to help decide which *will be best*, and under what conditions (Type B).[4]

NOTES REGARDING THE USE OF MATHEMATICAL STATISTICS

Prediction, mathematical statistics, and the prerequisite of statistical stability. As the operation of measuring a product is repeated over and over, whether on samples or the entire lot, a population consisting of a sequence of terms is generated. Each term represents a measurement, or a function of several measurements. The Type A problem is to find a method by which predictions concerning future terms of this sequence can be made with the highest possible degree of belief. This is the problem that the statistician is called upon to answer whenever the results of measurement and sample surveys form the basis of planning or other action. He is expected to answer it if possible, and to know when he can answer it. When he can, he is expected to give a better answer oftener than other people can give.

[3] See for example, H. F. Dodge and H. G. Romig, "Single sampling and double sampling inspection tables," *Bell System Technical Journal*, Vol. XX, January 1941, pp. 1–61.

[4] Page 31 of the 2d edition of Ezekiel's *Methods of Correlation Analysis* (John Wiley, 1941) is recommended at this point for an extension of these remarks.

In the state of randomness or statistical stability, the sequence will appear to have been drawn blindfolded and with replacement from a bowl of physically similar numbered chips. In this state, a distribution formed from terms of the sequence is *stable*, and it is then that predictions regarding future terms can be made by mathematical statistics with the highest attainable degree of belief. The Shewhart criterion[5] is valuable in assisting judgment regarding randomness.

When the criterion for randomness is not met sufficiently to warrant the application of mathematical statistics for predictions, the interpretation of the results of the survey or experiment must lean heavily upon the statistician's knowledge of the subject matter, and his ability to cooperate with experts.

The fundamental problem of mathematical statistics. The fundamental problem of mathematical statistics is to set *fixed limits* within which percentages of the next (e.g.) 1,000 terms of a random sequence will fall, and to set these limits efficiently from terms generated in the past. The problem is evidently one of prediction. The Shewhart methods were devised to aid judgment in deciding whether the sequence is sufficiently near random to permit the fundamental problem to be attempted, and as an aid in solving it when it can be solved. Wilks[6] has presented a pioneer piece of theoretical work on the subject. Confidence intervals and fiducial probability can not fully answer the purpose because the interval in the fundamental problem is a tolerance interval of specified width; it is not a random variable, but is *fixed*.[7] True, the random intervals in the theories of confidence intervals and fiducial probability become steady in large samples, but a large sample does not by itself exhibit evidence for or against stability (randomness) until it is broken down into rational subseries, as it will be by the application of the Shewhart methods. (Cf. the section, "Repeated patterns in subseries.")

As a simple illustration, let p be the ratio of white to black chips found in a sample that is drawn from a bowl containing white and black chips. When the sampling is random, with replacement, the average p will approach a statistical limit p_0 as the number of samples increases. The fundamental problem of mathematical statistics is to name the proportion of the next 1,000 samples for which p will lie within the

[5] Walter A. Shewhart, *The Economic Control of Quality of Manufactured Product* (Van Nostrand, 1931), Ch. XX.

[6] S. S. Wilks, "Determination of sample sizes for setting tolerance limits," *Annals of Mathematical Statistics*, March 1941, pp. 91–96.

[7] The distinction between the different kinds of prediction is illustrated on p. 50 and elsewhere in Shewhart's *Statistical Method from the Viewpoint of Quality Control* (The Graduate School, Department of Agriculture, 1939); also in fig. 11 of Deming and Birge's *Statistical Theory of Errors* (The Graduate School, Department of Agriculture, 1934, 1938).

interval $p_\infty \pm \sigma$. Whether p_∞ is the same as the ratio of white to black chips in the bowl is still another problem (next section).

Knowledge of the subject matter essential in the Type A problem. Let p' be the ratio of white to black chips in the bowl. It can be determined only by taking all the chips out of the bowl and examining each one, so in practical problems of sampling it remains unknown. Even when the sampling is carried out by a random operation, and the statistical limit p_∞ exists, it remains to decide whether the value of the unknown p' is anything like p_∞. To answer this question in real surveys the statistician must be concerned with the subject matter or even with psychology perhaps more than with mathematics. For instance, he must recognize the fact that the sampling and measuring may introduce biases, as for instance when the people in the sample behave differently (change characteristics) just because they in particular are under observation and the others are not. It will not do to issue one statement (prediction) as a statistician, and some other statement as an economist, population or social expert, agricultural expert, geologist, chemist, engineer, or anything else.

Sources of discrepancies in different surveys may arise from differences in definition and procedure: the auspices, frequencies of interrogation, training of the enumerators, supervisors, and editors, informants, pay and time allowed for the interviews, volumes of questions, way of asking the questions, and a host of other sources of discrepancy, will all affect the results. That large discrepancies can arise from small changes in the definitions and procedure is often not appreciated until a survey is repeated after a short time interval with as few changes in procedure as possible, or until it is compared with a similar one taken under different auspices at about the same time. If the surveys are carried out by sampling, there will of course be sampling errors to contend with, but sampling errors may be less troublesome than some of the other difficulties just mentioned. Increasing the samples to 100 per cent will eliminate the sampling errors, but not the discrepancies arising from differences in definition and procedure. Knowledge of the subject matter, and an appreciation of the limitations of measurement, are therefore necessary in the interpretation of results, whether they are obtained by complete counts or by samples.

Similar remarks apply to the physical sciences, wherein differences in the various instruments, procedures, definitions, states of wear and tear on the instruments, and the theoretical relations that are assumed to exist between the different quantities, play the counterpart of the differences in the procedures and definitions in the social sciences.

Knowledge of the subject matter essential in the Type B problem. The study of any Type B problem will require one or more investigations of Type A as part of the evidence for the action that is required. The necessity for knowledge of the subject matter in the interpretation of a Type A experiment still exists when the experiment is carried out as part of a Type B study.

Other evidence for a Type B prediction may come from studies of the underlying process from the standpoint of sociology, economics, mechanics, chemistry, geology, biology, or whatever may be involved. Such studies of course require considerable knowledge of the subject matter, but they may take the place of a vast amount of experimental data of Type A. Moreover, in the Type B problem it is a matter of judgment and knowledge of the subject matter to state the range of validity of a relationship, and to decide when enough situations have been covered to establish this validity with a sufficiently high degree of belief for the action required. Judgment is the result of scientific training—intuition if you like; but intuition, like the conscience, must be trained.

The importance of the design of experiments. Progress in a Type B study is enhanced if the Type A experiments necessary thereto are carried out with the greatest possible efficiency. The importance of the theory of complex experiments and other branches of mathematical statistics in the design of sample surveys and experiments can hardly be over-emphasized. However, it is one thing to design a sampling plan or experiment so that it *ought* to exhibit randomness and give results with a certain variance, but it is another matter to show that the results have the validity that was hoped for.

A word on the presentation of results. An analysis and evaluation of data, to show how closely the operation of sampling *satisfies the requirements intended,* ought never to be omitted in the presentation of the results of a Type A investigation. This analysis is part of the evidence, and is particularly important if mathematical statistics is used in the interpretation of the results. Without it, the reader does not know what confidence to place in the predictions that are made, and what action dare be recommended on the basis of them. Often when such analyses and evaluations have been made, it has been found that the sampling method did not work as intended; and what is more, reasons for spurious results are often found, with consequent improvement in interpretation. A sampling plan ought always to be designed so that the sample can be broken up into small samples, geographically or temporally or both—samples that may be too small for publication, yet

large enough to be compared with one another and examined for randomness and patterns (next section).

Repeated patterns in subseries. Significance. A pattern that is repeated under a wide variety of conditions may provide evidence for a degree of belief sufficient to justify a hazardous course of action in a Type B problem, even in the absence of a rational theory regarding the underlying causes. A repeated pattern may attain *scientific significance*, even though no one of the patterns by itself would seem worthy of note. Thus, if treatment C has invariably been found better than treatment D under a wide variety of soils and climates, favor toward treatment C would not rest so much on any "significance" calculated from either a single experiment or a combination of experiments, as on the apparent ability of treatment C to maintain superiority under any likely set of conditions. Other examples are afforded by numerous empirical laws. The necessity for breaking up a large sample and studying the patterns in small rational subseries (as in order of time) is in fact the kernel of the Shewhart methods,[6] which are as applicable in the Type B problems of the social sciences as they are in industry.

Occasionally, in an extreme case of a Type B problem, our knowledge of the underlying forces, however derived, may be so extensive that in view of the consequences of the action, it is sufficient to perform but one (Type A) survey or experiment, which will then constitute evidence for prediction with a high enough degree of belief to point to the course of action required. Under such circumstances, a single test of significance, for instance, or the calculation of a single confidence interval, may suffice for action. A single experiment, or a single test of significance, may also suffice in circumstances where the consequences of the action are not critical, even though knowledge of the underlying causes is not extensive: it may not be worth while to get more evidence. Of course, as was mentioned earlier, there are times when action must be taken without sufficient evidence.

EFFECT OF THE CLASSIFICATION ON THE PLANNING OF ENQUIRIES

Necessity for keeping in mind the ultimate objectives of an enquiry; are they Type A or Type B? Recognition of the two types of problems and methods calls for discrimination in the planning of any survey or experiment. Its aims must be kept in mind. What will the data be used for? Is the problem Type A or Type B? This classification will affect

[6] This point is also insisted on by Keynes: confer his *Probability* (Macmillan, 1929), pp. 407–08.

its design, the funds needed and how they should be spent, the amount of detail required in tabulation, the areas of tabulation, whether samples will best serve the purpose, and if so, what size of sample is required, and—most important—how the samples should be timed and distributed geographically. For example, where the action by definition or law depends on the state of a population as it exists on a certain date (as for allocation of funds by states, or conscription by age groups), the data required are purely for Type A purposes. To provide the necessary detail and accuracy it will be necessary to collect the data by a large enough sample, and if extreme detail is required, by a complete count. On the other hand, in a study of relationships, which would necessarily be a problem of Type B, it may be wise to carry out numerous small-sized experiments or surveys, spaced temporally and geographically in order to cover a wide variety of conditions (e.g., other city size groups, other climates, other soils), and thus obtain patterns in subseries. In the Type B problem, all experimental data, even complete counts, are but samples of what the underlying system of forces can and will produce.[9]

Some remarks on quality control in industry. Attaining control of quality during production is a Type B problem. The essential feature of the Shewhart statistical method[10] is to break up the inspection data into small rational subgroups in order of production, so that a large batch of data is studied, not simply as a large sample, but as an ordered sequence of small samples. Quick examination of the samples is made, and the results are plotted on a control chart so that *action can be taken at once* when the chart shows that something has gone wrong with the process of manufacture. Small samples suffice, because they are *not* for lot by lot evaluation (Type A), but for *control of the process* (Type B).

In the limiting state of stability, called statistical control or randomness, the formulas of distribution theory may be applied to the problem of determining what percentage of the inspection data from tomorrow's product will fall within certain *fixed limits*—the fundamental problem of mathematical statistics (q.v.). In this state the statistician may make full use of the calculus of probabilities, with no reference to further study of the process, because tomorrow's product is simply, in effect, some more product *drawn from the same bowl.* After a program of statistical control is in operation, disposition of the product, lot by

[9] These remarks are extended in a paper by W. Edwards Deming and Frederick F. Stephan "On the interpretation of censuses as samples," this JOURNAL, Vol. 36, March 1941, pp. 45–49.

[10] Cf. the section, "Repeated patterns in subseries. Significance;" also the reference to Keynes in footnote 8.

lot, becomes to greater or less extent a Type B problem instead of Type A, and the amount of inspection can usually be greatly diminished, with attendant savings in time, labor, and materials.

The state of statistical control is not easily attained. It has never been known to happen by intent alone. It is rather something to be achieved after weeks or months of effort. Every step accomplished toward the attainment of statistical control, however, results in savings in time and materials and cost of inspection, to both the producer and the consumer.

Frequent samples desirable in social and economic studies. Just as the examination of frequent small samples has accomplished so much in the Type B problems of industry, so also, the introduction of frequent small sample surveys of population and agriculture (perhaps 3 or 5 per cent) throughout the country would constitute a distinct advance in social and economic planning. They would provide a record of changing conditions, while these changes are taking place. Thus they would facilitate studies of the underlying cause systems that make the population what it is, and provide a better basis for planning than we have now. They would also enhance the value of the detail furnished by the complete census, and would give an indication of information that needs to be obtained in detail at the next complete census.

For like reasons, quick sample reports are supplementing or displacing the slower and unwieldy complete counts in many economic surveys needed in government Type B planning. The presence of some sampling error is often unimportant compared with the advantages of the quick and frequent returns made possible, usually at lower cost, by sampling. These advantages are appreciated when the distinction between the Type A and Type B problems is implicitly or explicitly recognized. In the time series of sampling results, it is usually the relation of one sample to the previous samples that provides a basis for action, rather than any one sample by itself.[31]

In this connection, it is easy to wax enthusiastic over the advantages that would accrue from a monthly report showing causes of death by various areas throughout the country. By the use of sampling, a report could be issued a few weeks after the close of each month. Epidemics could be recognized in their early stages, and steps taken at the right time to keep them from spreading. They could be traced as they move about. The complete annual report would still be valuable, but for other purposes, such as for detail by small areas, and for extremely

[31] See footnote 10.

rare diseases. This would be a control program as exciting as any in industry.[12]

ACKNOWLEDGMENT

Most of the material in this paper is scattered in other places, implicitly if not explicitly; nevertheless it has been found helpful in the work of the Census to bring it together in one place, and to put it into practice. In developing this presentation I have had the advantage of frequent and extensive discussions with Dr. Walter A. Shewhart and Mr. Harold F. Dodge of the Bell Telephone Laboratories. It is a pleasure to acknowledge the assistance of many friends, and to record the continued encouragement of Dr. Philip M. Hauser, Assistant Chief Statistician for Population in the Bureau of the Census. Courses given at the Graduate School of the Department of Agriculture, and lectures delivered at Georgetown University, Illinois, Cornell, Ohio State,[13] the Massachusetts Institute of Technology, Brown University, Catholic University, and Hunter College, have afforded me opportunities to expound these views and to hear the valued criticisms of students and many able statisticians. Presentations at the meeting of the American Statistical Association in New York in December 1941 were also helpful in this respect.

[12] This idea was suggested in a conversation with my friend Dr. Forrest E. Linder, Assistant Chief Statistician for Vital Statistics. The details for putting such a program into operation are now being studied.

[13] The lecture delivered at Columbus was published in the *Proceedings of the 1941 Ohio Conference of Statisticians*. It describes the two types of problems, with emphasis on the design of surveys in marketing research.

MEASURING NATIONAL INCOME AS AFFECTED BY THE WAR*

By Milton Gilbert
The Bureau of Foreign and Domestic Commerce

The subject "Measuring National Income as Affected by the War," is sufficiently broad to embrace many problems of income measurement. There are theoretical issues concerning the meaning of national income in wartime as either a production or a welfare measure that could be discussed. Such questions are being passed over, however, in favor of a problem of more immediate and practical interest—that of how expenditures for war purposes can be compared with national income so as to indicate what value of goods and services remain for civilian uses of various sorts. This emphasis is intended as an aid to those who have turned to national income as a practical tool in connection with responsibilities imposed by the war, and who have not the time to explore the subject in technical literature. Those thoroughly familiar with national income problems will find little that is new, apart from the estimates for recent years that are presented.

As is well known, the national income has risen to record levels under the stimulus of, what was formerly the defense program and is now, the war effort. (Department of Commerce concepts and estimates are used throughout this paper.) In 1941 it reached 94.7 billion dollars as compared with the 1940 total of 77.3 billion dollars, the gain of over 17 billions constituting the largest annual increase in our history. This rise in the national income is usually contrasted, in one form or another, with the change in the stimulus itself, that is, with the increase in armament outlays. To take the figure most commonly used, the total of defense expenditures and British (or Allied) armament purchases rose from about 4 billion in 1940 to about 15 billion in 1941, an increase in the neighborhood of 11 billion dollars.

Inasmuch as the national income is defined as the net value of the goods and services produced by the economic enterprises of the Nation, it would seem quite appropriate to make a direct comparison between these two aggregates. If that comparison is intended merely as a rough measure of the magnitude of the country's war effort or of changes in the size of the effort over time, not too much violence to the facts may be done. A real difficulty arises, however, when an attempt is made to draw inferences from the national income and war expenditures aggre-

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 29, 1941.

gates regarding changes in the level of civilian goods output. This difficulty may be highlighted by asking the following question: If 11 billion of the 1941 increase in national income was being utilized for national defense purposes, must there not have been very little increase in the value of goods and services utilized for civilian purposes? In fact, if the 6 billions remaining after the defense total is subtracted from the increase in output, is contrasted with the increase in national income due solely to the rise in prices, would there not seem to have been negligible change in the real output of consumption and capital goods for private purchase?

Anyone familiar with what happened to other reliable economic indicators during the past year would be likely to suspect that this conclusion is untenable and, consequently, that there is something wrong with the arithmetic producing such a result. This, of course, is the case. Defense expenditures and national income are not fully comparable aggregates. The one cannot be subtracted from the other and yield a significant remainder. Care must be exercised, also, in expressing war expenditures as a percentage of national income.[1]

This does not mean that either of these figures by itself has not a clearly defined meaning but merely that they represent different types of aggregate which are incomparable. To anticipate somewhat, the reason for this, briefly, is that defense expenditure is (firstly) a sum of *transactions* in the sense that it includes payments other than those for goods currently produced, and (secondly) a sum in which the goods and services included are valued at prices paid. The national income, on the other hand, is (firstly) a net value of current output in which the net is defined in a special way, and (secondly) a total in which the valuation of output is at costs paid or accruing to the factors of production rather than at sales prices to final users. Hence, in order to obtain a meaningful comparison between war goods and total output, adjustment is required in both war expenditures and national income. The character of these adjustments will be shown through an explanation of the various items in Table I. It should hardly be necessary to add that the estimates are preliminary, particularly those for 1941. The Bureau of Foreign and Domestic Commerce is engaged in making direct estimates of this sort in a thorough and careful way, refining and extending the outstanding work of Professor Kuznets in the field of commodity flow and capital formation, but this job is as yet incomplete.

One must begin a discussion of this sort by pointing out that there is

[1] See "Measuring the Economic Impact of Armament Expenditures," by R. W. Goldsmith, a paper presented at the 1941 meetings of the American Statistical Association.

no one correct measure of income or output that can be used indiscriminately in every type of economic problem. Any measure is correct so long as it is consistently worked out from a strictly defined concept, but the concept itself must be adapted to the analytical purpose in

TABLE 1

DERIVATION OF GROSS NATIONAL PRODUCT AT MARKET PRICES FROM NATIONAL INCOME*

(In billions of dollars)

	1939	1940	1941
National income	[illegible]	77.3	94.7
Plus: Corporation income, etc., taxes	1.5	2.8	6.9
Other business taxes	8.1	9.0	10.7
Accruing depreciation	6.4	6.5	7.0
Capital outlays charged to current expense	.6	1.0	1.8
Other business reserves	.8	.9	1.6
Inventory revaluation	.3	−.4	−3.2
Equals: Gross national product at market prices	[illegible]	97.1	119.5

* Since this paper was written estimates for the years 1929–1941 have been published in the Survey of Current Business, May 1942. Reference may be made to that publication for technical notes on the estimates.

Source: Bureau of Foreign and Domestic Commerce.

view. Problems of taxable capacity, of economic welfare, or of productivity all require different measures, even though it is considered desirable to call only one concept "the national income" just to avoid confusion. The same thing can be said about comparisons of war expenditures and total output. The concept of total product used must be framed according to the purpose for which the comparison is made. A definite objective has, therefore, been set here; the derivation of the total of consumers' purchases of market goods and services by the subtraction of war expenditures and any other non-consumer spending from total product. A more general usefulness may be served by the discussion in pointing out some of the statistical and conceptual difficulties in this type of estimating.

As already indicated, the national income cannot properly be used as the measure of output from which to deduct war expenditures and other non-consumer purchases so as to yield consumers' purchases, even neglecting for the moment the incomparability of the war expenditures total. The reason for this may be clarified by considering what measure of total output would be appropriate. Now, war expenditures are largely made up of two sorts of purchases of current output. In the first place, there are the purchases made by government of goods and services produced by private industry. Consequently, the measure of total output appropriate to the present purpose must contain the

value of the output of private enterprise at final market prices. This might be obtained by adding up the sales of each business unit, adjusting for its change in inventory, and then deducting its purchases from other business units. Net sales of all business summated would yield the desired value of product for private enterprise. This figure may best be thought of as the income from sales that would be shown by a consolidated income statement for all private enterprises, with adjustment for changes in inventory holdings.

In the second place, war expenditures are utilized to pay for goods and services produced directly by government; consequently, the measure of aggregate output must include the cost value of government production. This can be obtained by summing the payments made to factors of production employed directly by government. When this sum is added to the previous total of the value of private enterprise output, a total output aggregate would be had from which could be subtracted the various categories of non-consumer expenditures.

The national income, on the other hand, differs quite considerably from this measure of gross product. It is made up of the sum of the returns paid to or accruing to the various factors of production. It contains already, therefore, the cost value of government production specified above. In the sphere of private enterprises, however, while the national income contains the preponderance of the unduplicated charges against gross revenue as specified above, that is returns to factors, it does not contain all such charges and hence is not equal to gross revenues. The additions made to national income in Table I are designed essentially to arrive at consolidated revenues from sales by securing total charges against revenues. It should be pointed out that, since the data are not complete, one must be content with building up a total which falls a little short of consolidated gross revenues. As there is no reason to suspect any consistent bias in this difference, however, it can be assumed that the total obtained moves parallel to gross revenues.

The principal charges against gross revenues, in addition to returns to factors (wages and salaries, net rents, interest, dividends, and undistributed profits), are business taxes and accounting depreciation and depletion. Further, it is appropriate to include also business charges for special reserves and for bad debt losses. Consideration may now be given to why these adjustments to national income are either necessary or desirable, starting with business taxes.

The national income is defined as the net value of the economic goods produced. In giving specific content to this definition, the value

of output is taken net of several possible elements of duplication, two of which result in making the value too low for our present purpose. It is, of course, net of the value of the intermediate products of private enterprise used in the production process. Insofar as this means not counting the value of the steel, the textiles, and the chromium in addition to the final value of the automobile, this is of no concern since the result achieved is the same as that obtained from the gross revenues of a consolidated income statement.

But this idea of netness is carried a step farther to exclude the intermediate products of government, that is, the government services rendered to business. Business taxes may be viewed as not included in the national income, in part to avoid this duplication. In other words, since government services to business are already in the national income when government output is taken at cost, they may be considered as eliminated in measuring the net output of private enterprise by deducting part of the taxes paid by business from consolidated gross revenues of business.

If war expenditures are to be deducted, however, the measure of output cannot exclude the intermediate products of government, for the costs of prosecuting the war are partly a service to business. Should intermediate products of government be omitted from the measure of aggregate output, what would be left after deducting war expenditures might be something less than the total amount of remaining final products. Similarly, the amount of dollars paid by consumers for the output of private enterprise is gross of any services government renders to business and included in the product purchased. It may be seen that the issue hangs not on any implicit classification of government services but on the fact that the services to business are measured by some part of business taxes.

The national income is also net of, what for some purposes might be, the duplication of values involved in counting both the market value of privately produced goods inflated by taxes on those goods, and the government services rendered with the use of those taxes. So long as sales, excise, or other taxes are levied which raise prices above factor costs, this sort of duplication[1] is present, even though no government services are assumed to be rendered to business. In part, therefore, business taxes may be viewed as not included in national income to eliminate this duplication in the value of final output. Because this is done, the national income is strictly a measure of the value of net out-

[1] "Duplication" is not a good word in this connection, since for some purposes the inclusion of taxes is not "double counting."

put at factor costs, rather than at market prices. While each of these measures has its uses, the latter is absolutely required here. This is so both because the consumer buys goods at market prices and because the privately produced goods purchased with war expenditures are bought at market prices. Thus, as neither of the possible reasons for leaving business taxes out of the national income applies to the concept of output value required here, all business taxes are added to national income in Table I.

The term "business taxes" as used in this connection has nothing to do with incidence. It represents merely such taxes as are paid by or through business as a matter of administration, regardless of whether they are passed on in the form of higher prices or not.

The estimates of business taxes shown in Table I have two major components. The first includes corporate income, excess profits, and capital stock taxes. Taxes on personal incomes from unincorporated business are not added since they are not excluded from the national income. The second component is all other taxes paid by business to government units, with the exception of the pay roll taxes paid by employers under the Social Security system which are already included in the national income estimates under "Other Labor Income." All taxes are taken on an accrual basis since the profit estimates currently included in the national income are derived from business income statements in which profits generally are shown net of accrued taxes rather than tax payments. There may be exceptions to this procedure, of course, but hardly for tax liabilities which vary significantly from year to year or for the taxes affected by the substantial rate increases since the start of the rearmament program.

The next item added to national income in Table I is business charges for depreciation and depletion. Obviously, the purpose here calls for adjusting national income as estimated in the Department of Commerce with the actual accounting charges of corporations and with analogous estimates, consistent with the character of the profits estimates, for non-corporate business. If depreciation and depletion charges were constant through time, this adjustment would be unnecessary, provided that interest were centered in the changes in consumers' expenditures rather than in their absolute amount. The changes in the residual, after the other appropriate adjustments and aside from statistical errors, would be the changes in the value of consumers' expenditures. Depreciation charges today, however, have lost much of their stability; a direct comparison of national product at market prices and defense expenditures for this purpose is, therefore, an unwarranted

procedure which is improperly dignified by calling it "statistical." Today, several factors are making for much higher charges for depreciation and depletion. These are the change in the depreciation rates to be allowed by the Bureau of Internal Revenue whereby defense facilities may be amortized in 5 years, the recent and prevailing high level of capital formation on private account, and the accelerated rate of mining output.

Furthermore, it is thought desirable to use gross national product for comparison with war expenditures, rather than net national product, if only to emphasize that accounting depreciation charges constitute a very inadequate measure of capital consumption. This is particularly true in time of war when the unavailability of many types of new equipment necessarily means that obsolescence of old equipment is slowed down considerably. Nonetheless, comparison of war expenditures with net national product has its uses and is in no sense incorrect. It can serve, for example, to bring into focus the fact that net capital consumption is an important source of war finance in real terms.

Addition of "capital outlays charged to current expense" is desirable for the same reasons as depreciation and depletion charges. By so doing the concept of gross national product is made consistent with gross capital formation defined as all investment goods having an average life of three years or more. Both concepts are, therefore, made to conform with economic notions of gross investment and freed from the vagaries of accounting practice.

The "other charges and reserves" that have been added to national income in Table 1 contain special emergency and contingency reserves and charges for bad debt losses. The special reserves being set up by many business concerns because of the uncertainties of the present situation must be added because, like taxes, they are covered by sales and yet not in the current estimates of returns to factors of production. It should, perhaps, be made clear that inclusion of this item raises no question about the necessity for the setting up of unusual reserves by business management. Many of them are intended to cover anticipated losses of foreign assets and may well be too low for the loss eventually incurred. Neither capital gains nor losses, however, are included in the national income, nor are such charges relevant in deriving a total of consumers' purchases.

So far as charges for bad debt losses are concerned, to the extent that these represent consumer bad debts, goods and services of equivalent value actually do reach consumers. They must be added therefore to make possible the derivation of the estimate of the sales value of goods

passing into the hands of consumers. As for business bad debt charges, in *the consolidation of the accounts of business enterprises* it is the revenues of sellers gross of charges for business bad debt losses which cancel against the purchases of business buyers.[3] Hence it is necessary to add all bad debt charges to business profits which have been computed net of such charges, as well as the other unduplicated items used here, to build up the approximation of business consolidated gross revenues.

The adjustment for revaluation of inventory, also included in Table I, is of a different character than the preceding additions to national income. It is not a necessary adjustment, providing that the estimate of gross private capital formation subtracted later in reducing the value of output total to a residual of consumers' purchases contains the change in the book value of inventory rather than the current value of the quantity change in inventory. Since many persons are accustomed to working with historical inventory data which are adjusted for revaluation, as estimated by Professor Kuznets, their inclusion here only serves to emphasize the obvious need for using a current estimate that is on the same basis as historical data, whether these be inventory or gross national product series.

With all these items added to the national income, it may be seen that there is now an increase of more than 22 billion dollars from 1940 to 1941 to be distributed among the various types of expenditure as against the 17 billion increase in the national income. And this is after the inclusion of a negative inventory revaluation estimate of over 3 billion which significantly affects the distribution of expenditures. This figure, which amounted to nearly 120 billion in 1941, has been labelled, somewhat hesitantly, "gross national product at market prices," in the hope that the last three words will clearly distinguish it from Professor Kuznets' gross national product concept. It is certainly a grosser "gross national product" than that which has become familiar through his work.

It is now possible to proceed with the second part of the problem, subtracting the various sorts of non-consumer purchases from current gross output so as to leave the desired residual of consumers' purchases. This is shown in Table II.

There are only two points to be made in this connection. The first is

[3] The above statement may perhaps be insufficient to confirm the necessity of adding back charges for losses on business bad debts. The following example may be helpful. Firm A, selling only to B, computes its profits net of bad debt losses. Firm B, buying from A, computes its profits on the basis of contract prices for all purchases, not on the basis of the lower value, less by defaults, actually paid for purchases. Hence, the profits of A plus the profits of B plus all other consolidated charges against revenues except bad debt charges fall short of the revenues of B by the amount of bad debt charges.

that the [illegible] have [illegible] the value of goods taken [illegible] of current output and not [illegible] transactions not influencing current output. The [illegible] and compari[illegible] of [illegible] have to be avoided. From th[illegible]

TABLE [illegible]

[illegible] AT MARKET PRICES
([illegible])

[illegible]	[illegible]	1940	1941
[illegible]	[illegible]	97.1	119.5
[illegible]	1.4	2.8	13.5
[illegible]			−2.3
[illegible]	5.4	5.2	5.2
Total	6.3	8.3	8.2
[illegible]	95.0	10.3	21.6
[illegible]	77.0	80.8	94.9
[illegible]	2.0	2.3	2.7
[illegible]	.9	1.1	1.4
[illegible]	.9	1.0	1.1
Net [illegible] of goods and services	5.0	6.6	8.9
Net [illegible]	.3	1.4	.0
Net change [illegible]	−.2	−1.1	−.0
Net change [illegible]	.8	1.8	3.6
Total	3.5	4.5	1.1
[illegible]	11.0	11.6	19.1
Less [illegible]	72.0	68.2	75.8
[illegible]	7.0	8.3	10.3
[illegible]	[illegible]	57.9	65.5

* Revised

Source: [illegible]

prevalence of errors on both these scores, some discussion may be justified.

The first item deducted in Table II is national defense expenditures. Used here is the total as shown in the Daily Treasury Statement plus changes in the assets of national defense corporations except for changes in their cash balances and transfers from other government agencies. It will be noted immediately that this total differs considerably from the 15 billion dollar figure cited earlier because it excludes armaments purchased in this country by foreign governments. This is the proper procedure for a source of expenditure breakdown since foreign purchases appear farther down in the table under net change in investment abroad. Those thinking in terms of a type of product breakdown should keep in mind, obviously, that a higher total for war goods would affect a reduction in the net change in investment abroad and not in consumers' goods available. The same holds true for any shift from foreign

purchasing of armaments to lend-lease shipments—the latter being included in war expenditures but not in net change in claims abroad.

As it stands, however, the national defense expenditures total cannot be deducted from gross national product, because it is a sum of payments made rather than a sum of purchases of goods and services out of current output. The adjustment shown is introduced to make the two figures comparable. It contains, for example, a substantial amount of advance payments made to manufacturers holding defense contracts for which no goods have as yet been received. The adjustment for prepayments should be net rather than gross, that is, the net outstanding payments on which deliveries have not been made as of the end of the period in question. At the present stage in the war production program, of course, the net prepayments are still a large positive sum which, therefore, appears as a negative adjustment to defense expenditures.

The other principal items of expenditure which have no counterpart in current output in the United States and which consequently must be included in a negative adjustment are purchases of land or other existing assets, except those affecting the inventory estimate, and off-shore expenditure for either labor or materials. One might mention, in addition, checking accounts set up regionally to facilitate the operations of the Quartermasters Corps and minor intergovernment transfers. All these adjustments together make quite a substantial sum which can lead to significant error when a total of war expenditures is compared with or deducted from current output.

Because all business taxes were included in the computation of the value of gross output, total non-defense government expenditures for current output are deducted in the next two items. As with net defense expenditures, these are substantially different from a total of non-defense expenditures by the Federal Government and expenditures of state and local governments. Budgeted expenditures have been adjusted to eliminate such payments as intergovernmental transfers, direct relief, Social Security benefits, veterans' pensions, purchases of land, etc., as none of these appears in the estimate of gross national product. It may be mentioned that the output of public service enterprises, such as the post office or publicly-owned utilities, operating outside government budgets are automatically excluded here and, therefore, appear below under the total of private goods for private use.

Little comment is needed on the various categories of capital formation shown in the table except to emphasize again that they are gross expenditures and that they are strictly on private account. The inventory estimate is the book value change adjusted for revaluations, since

this adjustment was made in building up the gross national product. It represents, that is, the current year market value of the increase in the physical stock of inventories.[4]

With all these subtractions from gross product, there is obtained as a residual the measure of consumers' purchases which has been the objective of this calculation. It shows that there was a very substantial increase in consumers' purchases in 1941. The estimate could be adjusted for the increase in the cost of living to indicate how much more real

TABLE III
NATIONAL INCOME BY USE OF FUNDS
(Billions of dollars)

	1939	1940	1941
National income	70.8	77.3	94.7
Plus, Transfer payments from government	2.5	2.7	2.4
Less: Corporate savings	.4	1.3	2.6
Employment taxes	2.0	2.2	2.4
Direct personal taxes	2.9	3.0	3.8
Federal	1.2	1.3	2.1
State and local	1.7	1.7	1.7
Equals: Disposable income of individuals	68.0	73.5	88.3
Less, Consumer expenditures for goods and services	62.0	66.2	75.8
Equals: Net savings of individuals*	6.0	7.3	12.5

* Residual.
Source: Bureau of Foreign and Domestic Commerce

goods and services the consumer secured. This would not be the equivalent of consumption of individuals, however, for there are items of such consumption included in the government purchases, the most important of which to remember at this time being the food, clothing, and shelter provided to the armed forces. If one chooses to use the consumption of individuals as a rough measure of changes in material welfare, therefore, the increase in consumption provided out of government funds should be included.

A third table has been added in order to indicate the magnitude of personal savings implicit in the consumers' purchases and national income totals. This filling out of the picture is useful in analysis of the fiscal problem. Little explanation seems required except, perhaps, to mention that the estimate of personal taxes is taken on a payments basis. This is of some importance in any speculation about changes in the propensity to consume.

Before concluding it may be mentioned that the comparison of war expenditures and national product (either net or gross) discussed above

[4] Certain implications of the estimates for the year ahead are discussed by the writer in "War Expenditures and National Production," *Survey of Current Business*, March 1942.

is only one type of comparison that can be made. It is the useful type when the objective is the tracing of expenditure flows. If, however, one is interested in the disposition of economic resources, as must be the case in problems concerning the war potential, it is necessary to make the comparison of war and non-war output in terms of factor costs.

It may be seen from the previous discussion that it was inappropriate to compare national income with war expenditures directly because the national income was, in a sense, too net a figure for the war expenditures total—even apart from adjustment suggested for war expenditures. The process of converting national income to gross national product, therefore, was essentially one of increasing the size of the national product concept to make it fit the concept implicit in the war expenditures. Now, it is possible to achieve this comparability the other way around; that is, by reducing the war expenditures figure until it is just as net as the national income so that both aggregates are in terms of factor costs.

In order to do this one must allocate business taxes between war expenditures and all non-war expenditures, and then reduce the war expenditures total by the amount of business taxes associated with it. By this means national income and war expenditures would be rendered directly comparable, due account being taken of the other adjustments previously mentioned. Comparison of this type is implicit in the tables on net national income and net national expenditure contained in the British white paper on the Sources of War Finance,[5] and the allocation of taxes is made directly by Mr. N. Kaldor in his comments on the white paper in the *Economic Journal* for June–September 1941.

It should be emphasized that this way of handling the problem is not a mere difference in methodology; the results achieved can serve different purposes. Specifically, if the objective is an estimate of the real resources in terms of factor costs being devoted to the war effort as against those being utilized for civilian purposes, it is essential that the taxes implicit in the two categories of expenditures be eliminated. Similarly calculation of the war potential in terms of real resources must be made ex business taxes. The reason for this, obviously, is that there can be no presumption that factor costs are proportionate to market prices.

A few words of caution must be added concerning this use of the national income as a measure of the quantity of real resources currently utilized. It is subject to severe technical limitations. In the first place,

[5] *An Analysis of the Sources of War Finance and an Estimate of the National Income and Expenditure in 1938 and 1940*, Cmd. 6261, H. M. Stationery Office, 1941.

modern cost theory does not lend itself very readily to any such concept as the quantity of real resources. Furthermore, there is the difference in accounting methods as between government and private business, which accounting records form the basis of estimates of factor costs. Moreover, a complication is introduced by the fact that certain important elements of income get fixed outside of the market in time of war, for example, wages of draftees. Perhaps most of all, it is rather difficult in practice to make proper allowance for the fact that perfect competition does not rule and that returns to the factors of production are far from equal in all industries. Citation of these difficulties is not intended as a ban on the use of national income estimates in economic planning for war, but only as a contribution toward the best use of a highly valuable tool of analysis.

SAMPLING THEORY WHEN THE SAMPLING-UNITS ARE OF UNEQUAL SIZES*

By W. G. Cochran
Iowa State College

In sampling, the sampling-units are usually chosen so as to be similar in size and structure. With some types of population, however, it is convenient or necessary to use sampling-units that differ in size. Thus the farm is often the sampling-unit for collecting agricultural data, though farms in the same county may vary in land acreage from a few acres to over 1,000 acres. Similarly, when obtaining information about sales or prices, the sampling-unit may be a dealer or store, these ranging from small to large concerns.

In such cases the question arises: Should differences between the sizes of the sampling-units be ignored or taken into account in selecting the sample and in making estimates from the results of the sample? This paper contains a preliminary discussion of the problem, though further research is needed, many of the results given below being only large-sample approximations. It is convenient to consider first the problem of estimation, since it appears that the best method of distributing the sample depends on the process of estimation that is to be used.

THE PROBLEM OF ESTIMATION

To state the problem of estimation in mathematical terms, we assume that sampling units are drawn at random without regard to their sizes, and consider how to estimate the population total of some quantity y which can be measured on each sampling-unit. Associated with each sampling-unit is also a quantity x, which is called its *area* rather than its *size*, to avoid possible confusion between the terms "size of sample" and "size of sampling-unit." Some knowledge is assumed to be available about the values of x in the sample, and possibly also in the population.[1] In order to apply results from the statistical theory of estimation, it is also assumed that the number of sampling-units in the population may be considered infinite. Formulae applicable to the

* A paper presented at the 103rd Annual Meeting of the American Statistical Association in joint session with the Institute of Mathematical Statistics, New York, December 30, 1941.

Journal paper No. J-989 of the Iowa Agricultural Experiment Station, Ames, Iowa, Project No. 611.

[1] For some populations an alternative method of specification may be more appropriate. For instance, each sampling may consist of an integral number of sub-units, as in the case of human populations where the sampling-unit is a household and the sub-unit is a single person. The specification may be made in terms of the value of y per sub-unit and the number n of sub-units per sampling-unit (cf. Hanson and Hurwitz, 1942). Since this approach would not apply in the examples given at the beginning of this paper, it will not be considered here.

practical situation of sampling from finite populations can be obtained by adding suitable correction terms.

Stated in this way, the problem of estimation is a familiar one in mathematical statistics. If the joint frequency distribution of x and y in the population is known, the theory of estimation provides a routine technique leading to an efficient estimate of the population total of y and using to best advantage any available information about x. There are, however, difficulties in utilizing this method of approach. The joint frequency distribution is often known at best only vaguely from the available data, and may not appear to follow any of the few types of bivariate frequency distribution that have been studied. Further, there are strong administrative arguments for keeping the computations involved in making the estimate as simple as possible; these requirements may impose a bar on the use of estimates which, while highly efficient statistically, are rather difficult to compute.

Both difficulties can be met to some extent by restricting the estimates to those derived from the regression of y on x. For the calculation of regression equations, it is not necessary to describe completely the joint frequency distribution of x and y: we need only know how the mean value and the variance of y change as x changes. These can be examined from a graph or two-way table of the pairs of values of x and y constructed from any available data. If the form of the regression line and the relative weights assigned to different values of y are correct, the regression estimate is a *best unbiased linear estimate* as defined by David and Neyman (1938), though it is not a maximum likelihood estimate unless in addition the values of y are normally distributed within arrays in which x is fixed. The computations required for the simpler types of regression line are well-known and not unduly laborious.

ESTIMATES DERIVED FROM LINEAR REGRESSION

In the following sections it will be assumed that the quantity to be estimated is the population total of y; any formulae can easily be altered so as to refer to the estimation of the population mean per sampling-unit.

The simplest case occurs when the mean value of y is linearly related to the area of the sampling-unit, with constant variance; i.e. y is of the form $\alpha+\beta x+e$, where e has mean value zero and constant variance in arrays in which x is fixed. In this case, the linear regression estimate Y_1 for the population total of y is

$$Y_1 = N\{\bar{y}_s + b(\bar{x}_p - \bar{x}_s)\} \tag{1}$$

where N is the number of sampling-units in the population, b is the sample regression coefficient $S(y-\bar{y}_s)(x-\bar{x}_s)/S(x-\bar{x}_s)^2$, and the suffixes p and s refer to the population and sample respectively. It will be noted that this estimate requires a knowledge both of the total number N of sampling-units and of the mean value of x in the population.

In samples in which the x's remain fixed, the sampling variance of Y_l is

$$V(Y_l) = N^2\sigma_y^2(1-\rho^2)\left\{\frac{1}{n} + \frac{(\bar{x}_p - \bar{x}_s)^2}{S(x-\bar{x}_s)^2}\right\} \tag{2}$$

n being the number of sampling-units in the sample and ρ the correlation coefficient between y and x. The distribution of Y_l tends to normality as n increases, being exactly normal for any size of sample if y is normally distributed for fixed x. A sample estimate of this variance is obtained by substituting for $\sigma_y^2(1-\rho^2)$ the mean square s_d^2 of deviations from the sample regression line.

For comparison with other estimates we may require the average variance of the regression estimate under random sampling. From (2), this clearly depends on the form of the frequency distribution of the areas. Since the areas are essentially positive, their distribution will not in general be normal, except perhaps as an approximation. The mean value of (2) may be expanded in a series of inverse powers of n, the sample size. Retaining the two leading terms, we obtain

$$\bar{V}(Y_l) = \frac{N^2\sigma_y^2(1-\rho^2)}{n}\left\{1 + \frac{1}{n} + \frac{3+2\gamma_1^2}{n^2}\right\} \tag{3}$$

where γ_1 is Fisher's (1941) measure of relative skewness ($\gamma_1^2 = \kappa_3^2/\kappa_2^3$). If the areas were normally distributed, γ_1 would of course be zero, and the exact value for the term in curled brackets would be $(n-2)/(n-3)$, which agrees with the value given above to this order of approximation. With large samples the factor is close to unity.

In many problems the true regression line must pass through the origin, as for example when y represents corn acreage and x farm acreage. Even in such cases, it may be advisable to use the preceding type of regression, if it appears on examination that a straight-line regression not passing through the origin will provide a satisfactory fit, whereas it would be necessary to use a curvilinear regression in order to include the origin. If a straight line through the origin can be used, y being of the form $(\beta x + e)$, with constant residual variance, the regression estimate Y_o ($_o$ for origin) of the population total is

$$Y_o = N\frac{S(xy)}{S(x^2)}\bar{x}_p = \{\Sigma(x)\}\frac{S(xy)}{S(x^2)} \tag{4}$$

where $\Sigma(x)$ is the population total of the areas. The variance of Y_o is

$$V(Y_o) = \{\Sigma(x)\}^2\sigma_y^2(1-\rho^2)/S(x^2). \tag{5}$$

The number of sampling-units in the population does not enter into either of these formulae, which require only the population total of the areas.

The expression for the average value of this variance, under repeated random sampling, is rather complicated. If the distribution of the areas is not far from normal, the leading terms give

$$V(Y_o) = \frac{N^2\sigma_y^2(1-\rho^2)}{n(1+c_x)}\left\{1+\frac{2c_x(2+c_x)}{n(1+c_x)^2}\right\} \tag{6}$$

$c_x = \sigma_x^2/\bar{x}_p^2$ being the square of the coefficient of variation of x.

From formulae (3) and (6), we may compare the sampling errors of Y_1 and Y_o with that of the estimate Y_s (s for sampling-unit) which is obtained by multiplying the sample mean per sampling-unit by the total number of sampling-units, and is commonly used where sampling-units are equal in size. Since the variance of Y_s is $N^2\sigma_y^2/n$, the ratios of the three pairs of variances in large samples are as follows:

$$\frac{V(Y_1)}{V(Y_s)} = (1-\rho^2); \quad \frac{V(Y_o)}{V(Y_s)} = \frac{(1-\rho^2)}{(1+c_x)}; \quad \frac{V(Y_o)}{V(Y_1)} = \frac{1}{(1+c_x)}. \tag{7}$$

The additional factors involving $1/n$ and $1/n^2$ have been omitted from these expressions; they should be included in practical applications unless they are negligible. In large samples, both regression estimates are more accurate than the sample-mean estimate, the gain in accuracy being considerable if ρ is high. As would be expected, Y_o is more accurate than Y_1 when the true regression line is straight and passes through the origin, the increase in accuracy depending on the coefficient of variation of the areas.

These results must be interpreted with care. They indicate that in large samples Y_1 can never be less accurate, on the average, than Y_s. This statement was proved under the assumption that the true regression is linear (whether it passes through the origin or not); in the following section it will be shown to hold substantially even if the true regression is not linear. The conclusions about Y_o have a much more restricted validity, holding only if the true regression is linear and

passes through the origin. If the true regression line passes through the point $y=\alpha$ when x is zero, the estimate Y_e is biased, the bias tending, in large samples to the constant value $-N\alpha c_x/(1+c_x)$. Including this bias in the expression for the sampling error, we have, instead of (6)

$$V(Y_e) = \frac{N^2\alpha^2 c_x^2}{(1+c_x)^2} + \frac{N^2\sigma_y^2(1-\rho^2)}{n(1+c_x)} . \tag{8}$$

Since the component arising from the bias does not decrease as the sample size n increases, a sample size is always reached beyond which both Y_l and Y_r are more accurate than Y_e, unless α is zero. Thus Y_e cannot be recommended as an estimate unless it is known with considerable confidence that the true regression is straight and passes through the origin.

NON-LINEAR RELATIONS BETWEEN y AND x

It has already been pointed out that in many cases the investigator possesses only fragmentary knowledge of the true relation between the observations y and the areas of the sampling-units. Since a linear regression estimate may be used without any certainty that the population regression is linear, it is worth examining how Y_l is affected when the population regression is non-linear. Suppose that the true relation is of the form

$$y = \alpha + \beta x + \xi + e \tag{9}$$

where as before e is distributed with zero mean and unit variance, independently of x, and ξ is a non-linear function of x. For this reason, it may be assumed without loss of generality that ξ has zero mean and zero *linear* correlation with x. Following the usual algebraic development of linear regression, we find that the error of estimate

$$Y_l - \Sigma(y) = N\left\{(\bar{\xi}_s + \bar{e}_s) + (\bar{x}_p - \bar{x}_s)\frac{S(\xi+e)(x-\bar{x}_s)}{S(x-\bar{x}_s)^2}\right\} . \tag{10}$$

Taking the mean value over all possible samples of n sampling-units, all terms become zero except the second term in ξ, whose mean value does not vanish on account of the non-linear correlation between ξ and x. By a technique developed by Fisher (1929), this value can be expressed in terms of the semi-invariants κ_{ij} of the joint distribution of ξ and x, the first two terms of the series being

$$E\{Y_l - \Sigma(y)\} = \frac{N}{n}\left\{\frac{-\kappa_{12}}{\sigma_x^2} + \frac{1}{n}\left(\frac{2\kappa_{12}}{\sigma_x^2} + \frac{\kappa_{14}}{\sigma_x^4}\right)\right\} . \tag{11}$$

Thus the regression estimate is biased, the bias however tending to zero as the sample size is increased, since n appears in the denominator of (11). The numerator a_{12} of the largest term depends essentially on the correlation between ξ and x^2, i.e. on the quadratic component of the regression of y on x.

Formula (3) for the average sampling variance of Y_2 is also changed, but the change affects only the terms in $1/n$, $1/n^2$ etc. inside the curled brackets, the factor outside the bracket remaining $N^2\sigma_y^2\,(1-\rho^2)/n$, which in this case is equal to $N^2(\sigma_1^2+\sigma_\xi^2)/n$. Since the bias in Y_2 changes in inverse proportion to the sample size, while the standard error of Y_2 changes inversely as $\sqrt{n}$, the bias ultimately becomes negligible relative to the standard error if the sample is sufficiently large.

Thus, with samples large enough so that terms in $1/n$ are negligible, the ratio of the variance of Y_2 to that of the sample-mean estimate Y_1 remains $(1-\rho^2)$ even if the population regression is non-linear. This does not of course imply that Y_2 is an efficient estimate in this case. If the correct form of regression line could be fitted, the variance of the regression estimate would be reduced, in large samples, to $N^2\sigma_1^2/n$, as compared with $N^2(\sigma_1^2+\sigma_\xi^2)/n$ for Y_2. As would be expected, the relative loss of information with Y_2 depends on the ratio of the variance of the "non-linear" component ξ to the residual variance.

At least part of loss of accuracy could be recovered by adding terms in x^2, x^3, etc., to the regression, with a corresponding increase in the numerical computations. In order to use such regressions in constructing the estimate, however, additional population data about x are required. For a quadratic regression, for example, we must know both the population mean and variance of x to be able to calculate the regression adjustments to the sample mean $\bar{y}_s$. It is unlikely that these would be available without a complete frequency distribution of the population by area of sampling-unit. Where such complete information is available, there is an alternative method of estimation which will be discussed later.

It was previously remarked that when the population regression is linear, an unbiased sample estimate of the variance of Y_2 is obtained by substituting the residual mean square s_d^2 in place of $\sigma_y^2(1-\rho^2)$ in formula (2). With a non-linear population regression, s_d^2 is a biased estimate of $\sigma_y^2(1-\rho^2)$, but again the bias is inversely proportional to n, becoming negligible in large samples.

To summarize, with large samples the linear regression estimate is unbiased, and the standard formula gives an unbiased estimate of its variance even when the population regression is non-linear. There is,

however, a loss of efficiency which remains fixed in large samples. While no exact small-sample theory has been reached, it appears that both the estimate itself and the estimated variance are biased in small samples.

WEIGHTED REGRESSIONS

Thus far we have considered the case in which only the mean value of y changes as x changes. The variance of y may also change, particularly so if there is considerable variation in the areas of the sampling-units. The theory of regression has been extended to meet this case, provided that the ratios of the variances of different values of y are known exactly, a condition which rarely if ever holds in problems of this type. If the true residual variance of y_i is σ_i^2, and $w_i = 1/\sigma_i^2$ the best unbiased linear estimate Y_{wt} is

$$Y_{wt} = N\{\bar{y}_w + b_w(\bar{x}_p - \bar{x}_w)\} \tag{12}$$

where $\bar{y}_w = S(w_i y_i)/S(w_i)$, $\bar{x}_w = S(w_i x_i)/S(w_i)$ are weighted sample means and $b_w = Sw_i(x_i - \bar{x}_w)(y_i - \bar{y}_w)/Sw_i(x_i - \bar{x}_w)^2$ is the weighted sample regression coefficient. For a fixed set of x's the sampling variance of Y_{wt} is

$$V(Y_{wt}) = N^2\left\{\frac{1}{S(w_i)} + \frac{(\bar{x}_p - \bar{x}_w)^2}{Sw_i(x_i - \bar{x}_w)^2}\right\}. \tag{13}$$

It will be noticed in (12) that Y_{wt} remains unchanged if instead of the correct weights w_i we use numbers $w_i' = \lambda w_i$ which are proportional to the weights; i.e. only the *relative* weights assigned to different values of y need be known in order to calculate Y_{wt}. Formula (13) for the sampling variance cannot be used however unless the actual values of the weights are known. If only relative weights w_i' are known, an unbiased sample estimate of (13) is given by

$$S^2(Y_{wt}) = N^2\frac{Sw_i'(y_i - Y_i)^2}{(n-2)}\left\{\frac{1}{S(w_i')} + \frac{(\bar{x}_p - \bar{x}_w)^2}{Sw_i'(x - \bar{x}_w)^2}\right\} \tag{14}$$

where $Sw_i'(y_i - Y_i)^2/(n-2)$ is the weighted mean square of deviations from the sample regression, using w_i' as weights.

In practice, before these formulae can be used, it will be necessary to estimate the residual variances, and hence the weights, from the results of the sample and any other comparable data. Baker (1941) has recently discussed this problem for the case in which the x's fall into a number of distinct groups, all x's having the same value within each group. More generally, the x's will show a continuous range of varia-

tion. Space does not permit a detailed investigation of the best procedure for estimating the weights in this case. If the weights are presumed to change continuously as x changes, the first step seems clearly to subdivide the range of x into a number of groups. The residual variance of y within each group can then be estimated by fitting an unweighted linear regression of y on x separately for each group. From these results, the relation between the residual variance and the area can be studied, and a smooth curve drawn to give the variance as a function of x. The weight to be assigned to any value of y is then obtained by noting the area of the sampling unit, reading the curve, and taking the inverse of the variance.

The greater the number of groups, the more points are available for appraising the relation between variance and area. A further advantage of having many groups is that if the range of x is small within the groups, the within-group correlation between y and x may be negligible, so that the total within-group mean square of y may be used as equivalent to the residual mean square, thus obviating the necessity of fitting a regression within each group. However, as the grouping is made finer, the number of observations within each group decreases, leading to less accurate estimates of the within-group variances. The optimum number of groups is not clear without further examination, though at a guess it seems advisable to have at least 20 observations in each group.

The estimated weights are, of course, subject to sampling errors. These errors have two consequences. The estimate Y_{wt} is not as accurate as it could have been made if the true weights had been known. This loss of accuracy is unavoidable, the somewhat laborious process described above for estimating the weights being an attempt to reduce the loss to a minimum. Secondly, and somewhat more seriously, both formulae (13) and (14) give biased estimates of the sampling variance of Y_{wt} *even in large samples*, i.e. even ignoring the correction terms of order $1/n$ which have appeared in previous formulae. If w_i' are the estimated and w_i the true weights, the correct sampling variance of Y_{wt} in large samples appears to be

$$N^2 S\left(\frac{w_i'^2}{w_i}\right) \Big/ (Sw_i')^2. \tag{15}$$

Substituting w_i' for w_i, formula (13) gives for large samples

$$N^2/S(w_i') \tag{16}$$

while (14) gives, on the average

$$N^2 S\left(\frac{w_i'}{w_i}\right) \bigg/ nS(w_i'). \tag{17}$$

All three formulae agree if $w_i' = w_i$ for all i, this being the only case in which (16) gives the correct result, except by chance. However (17) also gives the correct result whenever $w_i' = \lambda w_i$ for any value of λ, and in general (17) is less subject to error than (16). Thus the process outlined above for estimating the weights should be regarded as leading merely to *relative* weights, formula (14) being used to estimate the sampling variance of Y_{wl}. If some idea can be formed of the probable magnitude of the errors in the estimated weights, formulae (15) and (17) can be compared to assess whether the estimated sampling error of Y_{wl} is likely to be greatly or only slightly biased.

Formulae (15) and (17) also agree if all the estimated weights w_i' are chosen equal, whether the true weights are equal or not. Thus, if we fit an *unweighted* regression, using Y_l as the estimate, the formula previously given for the estimated sampling variance of Y_l is still unbiased in large samples when the true weights vary. Considering how frequently the unweighted linear regression is used in statistical applications, it is reassuring to find that at least in large samples the standard formula for the estimated variance remains reliable even if the population regression is non-linear or if the true weights vary.

In view of the labor involved in estimating weights and fitting a weighted regression, it may sometimes be questioned whether the gain in accuracy is sufficient to compensate for the extra work, particularly so if the true weights do not appear to vary greatly. To obtain some idea of the gain in accuracy, we may note from either (15) or (17) that if an unweighted regression is used, the variance of Y_l is approximately $N^2S(1/w_i)/n^2$. From (13), the *maximum possible* accuracy attained by a weighted estimate is $N^2/S(w_i)$ to the same order of approximation. Thus the relative accuracy of Y_l to Y_{wl} cannot be less than approximately

$$\frac{V(Y_{wl})}{V(Y_l)} = \frac{n^2}{S(w_i)S\left(\dfrac{1}{w_i}\right)} = \frac{n^2}{S(\sigma_i^2)S\left(\dfrac{1}{\sigma_i^2}\right)}. \tag{18}$$

By inserting a series of values of σ_i^2 to represent the range of variation in a practical case, this formula gives some idea of the relative accuracy attained by an unweighted regression. If the true variances do not change greatly, a rough approximation to (18) is $1/(1+c_v)$, where c_v is the coefficient of variation of the variances. Thus, for example, if

the variance σ_i^2 appears proportional to the area of the sampling-unit, the relative accuracy is about $1/(1+c_x)$.

ESTIMATION OF THE MEAN PER UNIT AREA

One type of weighted regression leads to an estimate which is particularly simple to calculate, and has proved serviceable in estimating crop acreages in agricultural sampling. If a weighted regression passes through the origin, the corresponding estimate Y_{wo} is

$$Y_{wo} = \frac{S(w_i x_i y_i)}{S(w_i x_i^2)} \Sigma(x_i). \tag{19}$$

Suppose that the variance of y_i increases proportionally to the area x_i; in this case $w_i = 1/x_i$, and (19) reduces to

$$Y_a = \frac{S(y_i)}{S(x_i)} \Sigma(x_i). \tag{20}$$

Thus to calculate Y_a (a for area), the sample total of y is divided by the total area of all sampling-units in the sample, giving a mean per unit area, which is then multiplied by the total area in the population. From the conditions mentioned above, Y_a is a best unbiased linear estimate if the mean value and the variance of y both change proportionally to x.

Goldberg (1942) has studied the sampling distribution of Y_a for any type of joint frequency distribution of y and x. He has shown that in general Y_a is biased, the leading term in the bias being

$$\frac{N y_p}{n}(c_x - \rho\sqrt{c_x c_y}) \tag{21}$$

while the first approximation to the variance is

$$V(Y_a) = \frac{N^2 y_p^2}{n}(c_x + c_y - 2\rho\sqrt{c_x c_y}). \tag{22}$$

Thus in large samples the ratio of the bias to the standard deviation is proportional to $1/\sqrt{n}$. By examining the ratio as a function of ρ, it may be shown that the ratio cannot numerically exceed the coefficient of variation of x, divided by $\sqrt{n}$.

Since Y_s and Y_a are the two simplest estimates to calculate, it is of interest to compare their sampling variances. For samples sufficiently large so that (22) may be used as the variance of Y_a, Goldberg (1942) has shown that Y_a has a smaller variance than Y_s whenever $\rho > \frac{1}{2}\sqrt{c_x/c_y}$ and vice versa.

ESTIMATION BY USING POPULATION WEIGHTS

If a complete tabulation of the areas of all sampling-units in the population is available, the areas can be sub-divided into groups or strata, an estimate of the total of y being made for each stratum. While this procedure could be carried out with all the estimates previously discussed, this investigation will be confined to the simplest estimate Y_1. If $n_1, \ldots n_k, N_1, \ldots N_k$ are the numbers of sampling-units in the sample and population respectively for the k groups, the estimate Y_{g1} of the population total over all strata is

$$Y_{g1} = (N_1\bar{y}_1 + \cdots + N_k\bar{y}_k) \tag{23}$$

the sampling variance being

$$V(Y_{g1}) = \left(\frac{N_1^2\sigma_1^2}{n_1} + \frac{N_2^2\sigma_2^2}{n_2} + \cdots + \frac{N_k^2\sigma_k^2}{n_k}\right) \tag{24}$$

where $\sigma_1^2 \ldots \sigma_k^2$ are the within-strata variances of y.

As shown by Neyman (1934), this variance is smallest, for a fixed total size of sample, when the sample is distributed amongst the groups so that n_i is proportional to $N_i\sigma_i$. To retain comparability with previous estimates, however, we will assume that the sample is chosen at random.

In large samples, the average value of (24) works out approximately as

$$\bar{V}(Y_{g1}) = \frac{N^2}{n}\left\{\frac{(N_1\sigma_1^2 + \cdots + N_k\sigma_k^2)}{N}\left(1 - \frac{1}{n}\right) + \frac{k}{n}\,\frac{(\sigma_1^2 + \cdots + \sigma_k^2)}{k}\right\} \tag{25}$$

where n and N are as before the total numbers of sampling-units in the sample and population respectively. The expression inside the curled brackets contains both a weighted and an unweighted mean of the within-strata variances of y.

If all within-strata variances are the same, this reduces to

$$\bar{V}(Y_{g1}) = \frac{N^2\sigma^2}{n}\left(1 + \frac{k-1}{n}\right). \tag{26}$$

By increasing the number of x-groups with a given sample, the within-group variances are presumably decreased, since that portion of the variances of y which is due to variation in x is decreased by cutting

given the range of x within each group. However, the factor involving k is of course increased as k increases, so that a point is reached beyond which a further increase in the number of groups will result in less accuracy. From (26) it follows that in the case of equal variances the factor involving k is relatively unimportant provided that $(k-1)/n$ is less than say 5%, which holds if the average number of observations per group exceeds 20.

On comparing (26) with (14), Y_{st} is found to be somewhat less accurate than the linear regression estimate Y_l if the true population regression is linear with equal variances. This follows because the within-stratum variance σ^2 cannot be less than $\sigma_y^2(1-\rho^2)$, while the additional factor in $1/n$ is also larger for Y_{st} than for Y_l. This conclusion was to be expected, since under the conditions mentioned Y_l is a best unbiased linear estimate. If however the relation between y and x is markedly curvilinear or discontinuous, Y_{st} may be superior to Y_l, since the variation in y arising from any type of relation with x can be reduced by a suitable choice of strata, whereas Y_l eliminates only the effects of the linear component of the relationship. Moreover, Y_{st} is an unbiased estimate for any type of relation between y and x and any size of sample. Similarly, an unbiased estimate of the variance of Y_{st} is always obtained by substituting the sample within-strata mean squares in (24).

Similar comparisons can be made between Y_{st} and the weighted linear regression estimates by means of the formulae given for the sampling errors. Goldberg (1942) has discussed briefly the properties of Y_{sm}, the corresponding weighted estimate derived from the sample mean per unit area within each group.

FURTHER NOTES

Some apology is needed for presenting in the previous sections a number of large-sample approximations without guidance as to the limits within which these apply. Unfortunately these limits depend on the form of the joint frequency distribution of x and y, and could not be specified more definitely without a classification of the types of frequency distribution. Moreover, in extensive surveys, where problems of organization are difficult, biases may arise through the method of selecting the sample, incompleteness in the returns, and errors in reporting or recording the data. Such biases, while affecting the accuracy of the estimates, may not be measured by the formula for the sampling error, so that a rough approximation to the sampling error is often sufficient for practical purposes.

If the correct form of regression is used, population estimates derived from regressions remain unbiased in non-random sampling, provided that all sampling-units *with the same area* have an equal chance of selection. Thus the large sampling-units might be allotted a greater chance of inclusion in the sample, this procedure giving a more accurate estimate whenever the variance of y increases as x increases. On the other hand, if the method of selection discriminates in favor of certain sampling-units amongst those of the same area, bias may arise.

The formulae in this paper will of course apply to any variable x which is correlated with y. For example, in agricultural sampling, where the sampling-unit is sometimes a fixed area of land, x may be taken as the number of farms in the area, the total farm land or the total crop land, according to which gives the highest correlation with y.

In developing correction terms to be applied where an appreciable fraction of the population is sampled, the initial difficulty is that of defining a regression in a finite population. Writing $y=\alpha+\beta x+e$, we may suppose that e has no linear correlation with x in the finite population, but if we attempt to postulate that e is uncorrelated with any power of x, the number of conditions to be satisfied is greater than the number of values of e available, so that e and x cannot be independently distributed in the sense in which this term is applied with infinite populations. An alternative approach is to regard the finite population as a random sample from an infinite population in which e and x are independent. From a preliminary investigation and from Goldberg's (1942) work, it appears that the first approximation consists in multiplying formulae (2), (3) and (22) for the sampling-variances, and formulae (11) and (21) for the biases by $(N-n)/N$, this being the same correction as in the case of the sample-mean estimate Y_s. In formula (24), each term is multiplied by the corresponding factor $(N_i-n_i)/N_i$. For Y_o, the estimate derived from a straightline regression through the origin, and Y_{wl}, the weighted linear regression estimate, further investigation is needed. The difficulty arises because these two estimates do not equal the true population total when the sample consists of the whole finite population; i.e. they are *inconsistent* in the sense of Fisher (1941) whereas Y_s, Y_l, Y_q, and Y_{ql} are always *consistent.*

For sampling surveys in which the areas x of the sampling-units are unequal, the properties of various estimates of the population total of some observed quantity y are discussed, these estimates being mostly derived from the regression of y on x. In order of ease of calculation, the

estimates are as follows; Y_s, derived from the sample mean per sampling-unit; Y_a, derived from the sample mean per unit area; Y_{ps}, a weighted form of Y_s, using population weights; Y_c and Y_d, based on unweighted linear regressions; and Y_{wc} and Y_{wd}, using weighted regressions. The conditions under which each estimate is most efficient are described, with various comparisons of their relative efficiencies.

REFERENCES

Baker (1941) this JOURNAL, 36, 500.

David and Neyman (1938) *Statistical Research Memoirs*, II, 105.

Fisher (1929) *Proceedings of the London Mathematical Society*, 30, 199.

Fisher (1941) *Statistical Methods for Research Workers*, Edinburgh, Oliver and Boyd, 8th ed.

Goldberg (1942) (Not yet published).

Hanson and Hurwitz (1942) this JOURNAL, 37, 89.

Neyman (1934) *Journal of the Royal Statistical Society*, 97, 558.

THE IMPORTANCE OF HOSPITAL MORBIDITY DATA FOR THE COMMUNITY*

BY MARTA FRAENKEL
Welfare Council of New York City

OFFICIALS and others responsible for administrative planning have had the benefit of a periodic census giving a solid basis of demographic and economic data on the total population. In one vast and important field, however, they have been forced to plan without this basis of statistics. This field is the care of the sick—the provision of medical and nursing care and the provision of hospitals.

Need for information in this field has always been present, but today that need is especially urgent. In recent decades the scene has been changing: A successful fight against acute infectious disease has pushed this group of conditions into the background while an aging population has brought chronic diseases more and more into the foreground.

The volume of disease is large and diseases vary greatly in kind and in seriousness; the borderline between sickness and health is not always clear cut. The combination of these factors may be the reason that a general periodic census of disease by diagnosis has never been launched. Knowledge of the occurrence of specified diseases in the population as a whole has been restricted to that which can be gained through an analysis of the causes of death and through the current data on the communicable and occupational diseases that are reportable by law.

Mortality statistics, it is true, have been developed to a high plane but data on the causes of death permit only limited conclusions as to the total volume of sickness which calls for medical, nursing or hospital care of the patients.

In the past, special studies have produced a body of interesting and useful information. But these studies have been restricted in size or in scope. The National Health Survey, by its very size and character, cannot be considered as the cornerstone of a routine reporting system that might furnish the periodical data necessary to community plans for the care of the sick.

Another approach is necessary. Since the reporting, as a routine procedure, of all illness in the population appears to be too difficult both on account of its vastness and its inclusion of ill-defined minor conditions, the diagnostic data on hospitalized patients are suggested as a regular source of information, especially in large urban areas.

* A paper presented at the 103rd Annual Meeting of The American Statistical Association, New York, December 28, 1941.

It is admitted that data on hospitalized morbidity are a compromise. Is this compromise acceptable? This question may be answered in the affirmative if the limits of the information made available are permanently kept in mind. Data on hospitalized morbidity reflect an unknown proportion of the total volume of illness, but it is safe to assume that they would afford a fairly complete record of the major diseases. Most of the minor conditions will, however, be excluded. One might wonder whether their inclusion would really prove an asset. Might not the enormous volume of minor illnesses and the absence of precise definition of such illnesses confuse the picture?

The collection from hospitals of morbidity data classified by disease should not meet insurmountable difficulties. These data would have to be recorded and collected by each hospital in accordance with standard rules and then to be forwarded to a central collecting agency. Since, for many years, some information has been thus reported, the major innovation would be the inclusion of the patients' medical diagnoses, recorded in accordance with a suggested classified list of diagnoses.[1] In suggesting the collection of data on hospitalized patients, the sources of morbidity data to a known and not too high number is limited. The record rooms of the hospitals are professionally directed by librarians accustomed to similar work and no doubt able to take up the additional task in a competent way.

The volume of hospitalized sickness without specification by diagnosis has been known for some years: the municipal hospitals report to their central administration; the voluntary hospitals, if they provide care to patients who are public charges, have to report to public welfare departments, and if they participate in a community drive for the support of voluntary hospitals, to the fund raising agency. Only proprietary hospitals are exempt from the obligation of any central reporting.

The a priori assumption that the hospitals' own reports might serve as a source of information on diseases of patients receiving hospital care, is for the present not correct. Most of the annual reports focus not on the patients treated for specific diagnoses, but on the service rendered by the several departments of the institution. In New York City, where the Department of Hospitals maintains a well-developed statistical service, the reports of the municipal hospitals include data on diagnoses.

A demonstration project, the *Hospital Discharge Study*, was under-

[1] A Classified List of Diagnoses for Hospital Morbidity Reporting, Welfare Council Publications 1030, Volume IX, Research Bureau, Welfare Council of New York City.

taken by the Work Projects Administration[2] and the Research Bureau of the Welfare Council of New York City, to explore the possibilities of hospital morbidity data with specification by diagnosis and the feasibility of their collection. This study, made under the direction of Dr. Nova R. Deardorff, covered 576,623 discharges, during the year 1933, from 113 of the 134 hospitals then in existence in New York City.[3] The information secured through the tabulation and analysis of these records is used to illustrate some of the comments which follow.

The two main elements determining the need for hospital care in a community are variable; that is, the proportion of patients suffering from specific diseases is not constant and the need for hospital care of patients with these diseases is changing.

The changes in the disease picture of a community mainly result from the changing composition of the population: A decrease in the birth rate and an increase in the average age are the main factors, with occupational changes being a third. In 1900, 30.7 per cent of the population in New York City were under 15 years of age and 2.8 per cent over 65; in 1940 the corresponding figures were 19.8 and 5.5 per cent, respectively, according to the data of the Bureau of Vital Statistics of the New York City Department of Health.[4]

What does such a change in the age composition of the population mean in terms of disease? Certain communicable diseases, such as scarlet fever, measles and diphtheria, are recognized to be the classic childhood diseases. In addition, such conditions as, for instance, rheumatic fever, mastoiditis and otitis media, are most common among the younger age groups. The *Hospital Discharge Study*, referred to above, shows for New York City among the two age groups of early life (under 5 years and 5 to 14 years), 352 and 199 discharges with the diagnosis of mastoiditis and 356 and 79 with otitis media per 100,000 of the population in these age groups. These acute conditions have a very small incidence in old age, namely, 24 and 17, respectively, per 100,000 persons of 65 years and over.

Among hospital patients of 65 years of age and over there were per 100,000 population of this age group 1,049 discharges with the diagnosis of malignant neoplasm, 798 with cerebral hemorrhage, 2,535 with cardiac disorders and 2,110 with vascular conditions. These few examples

[2] *Work Projects Administration of the City of New York*, O. P. No. 65-1-97-21 W. P. 6.

[3] *Hospital Discharge Study*, An analysis of 570,023 patients discharged from hospitals in New York City in 1933, by Nova R. Deardorff, Ph.D., and Marta Fraenkel, M.D., Volume I, "Hospitals and Hospital Patients in New York City" (in press), Volume II, "Hospitalized Illness in New York City," and Volume III, "Reporting of Illness by Hospitals" (in preparation).

[4] *Quarterly Bulletin*, Department of Health, City of New York, Volume IX, No. 3, 1941.

may suffice to show the changes in the disease picture of a community which parallel changes in the age composition of the population. Current knowledge of the prevalence of diseases within given age groups is the first step for a proper preparation to meet the need for care.

The second inconstant factor in the picture is the changing need for hospital bed care for particular diseases. The demand for hospital beds may be affected by new therapeutic measures; the switch from surgical intervention to radiation in the treatment of malignant neoplasms of selected sites, for instance, has reduced the amount of hospital bed care needed by cancer patients. To forecast changes in the volume of hospital care needed, knowledge of the occurrence of the underlying diseases is a prerequisite.

Changes in the need for hospital bed care may also result from the adoption of new policies. Hospital bed care is only one method of caring for sick people. It is not always the most appropriate one; it is expensive and patients, especially old, chronically sick persons, are frequently unwilling to stay in the hospital for extended periods. Care in institutions such as homes for the chronically sick, homes for the infirm aged, or convalescent homes is more economic to the community and often is preferred by the patient. In many instances, institutional care of patients with chronic diseases or of convalescent patients can be prevented or curtailed if social work planning based on the knowledge of the existing demand provides an adequate supply of extra-institutional care. Visiting nurse services and visiting housekeeper services, supplemented by medical supervision, may be the tools used in such a program. But any planning for these services must be based upon knowledge of the occurrence and distribution of the underlying diseases.

What type of information can be expected from data on hospital morbidity? Five problems on which the study of hospital discharge data threw some light may be cited as examples.

First: What are the diseases found most frequently among hospital patients?

Obstetrical work, which is not a fight against disease but constructive health work, according to the data of the *Hospital Discharge Study*, accounts for a larger share of New York City hospital patients than any other condition. Among the 576,023 discharges, the 59,684 delivered women and the 58,507 live births amounted to 20.5 per cent of the total discharges.

Tonsil conditions with 68,384 discharges, or 11.9 per cent of the total, were the second largest single item. Approximately 17,100 cases, or 3.0 per cent of the total, were diagnosed as acute appendicitis; 16,008 cases,

or 2.8 per cent, as tuberculosis, and 10,302 cases, or 1.8 per cent, as inguinal hernia.

The second question is: What is the occurrence of various diseases when the data are analyzed by sex, age and color of the hospital patients in relation to the total population of the same sex, age and color?

If the number of female patients whose hospital stay was caused by childbirth are eliminated from the figures, male hospital patients form a slightly higher proportion of the total than did the females, namely, 53.8 per cent. This is caused by the high occurrence of preponderantly "male conditions": 08 per cent of the 21,302 patients with fractures, 79 per cent of the 13,424 with hernias and 86 per cent of the 15,570 with alcohol poisonings were male.

The importance of the age data has already been stressed above.

The distribution of the discharges by color shows 71 white hospital patients per 100,000 white population, but 126 Negro hospital patients for the corresponding population group. Tuberculosis and venereal diseases contribute to the excess of hospitalized Negro patients with 16.2, 31.8 (syphilis) and 22.1 (gonorrhea) per cent of the total cases respectively. Ulcer, scarlet fever and diabetes mellitus are some of the diseases for which Negroes are less frequently hospitalized; Negroes constitute 3.8, 4.8 and 5.0 per cent of the total cases respectively.

How greatly does the length of stay in a hospital vary with the particular disease? This third question is important to any planning. An estimate of the number of hospital beds needed must be based on the distribution of the days' care needed by patients with various conditions. Long as well as short hospital stays were caused by acute as well as by chronic conditions.

The data, moreover, show to what extent the amount of hospital care needed by patients with a given diagnosis depends upon economic factors. For some diseases, patients can be discharged at an earlier period of convalescence from voluntary hospitals than from municipal hospitals; for instance, 75.6 per cent of the acute appendicitis patients were discharged from voluntary hospitals after a stay of 8 to 14 days but only 44.2 per cent of those in municipal hospitals were discharged so early.

What is the mortality rate among hospital patients, is a fourth question.

The mortality in hospitals of patients covered by the *Hospital Discharge Study* was 5.6 per cent of the total. This rate for all diagnoses hides rates for specific diseases which range from 0.7 for obstetrical and some gynecological conditions to 49.8 per cent for septicemia and 60.9 per cent for peritonitis. The deaths in hospitals cover only 42.2 per cent

of the total cardiac deaths in the city but 100.0 per cent of the deaths caused by epidemic encephalitis.

Planning for adequate medical and nursing care should take into consideration the fact that patients with some diseases are transferred to hospital care in an extreme condition and that those with other diseases often stay in the hospital for weeks and months until death comes.

Fifth and last: Can hospital facilities be planned in accordance with the population of an area?

The planning for certain community facilities such as schools, churches and recreational facilities, is based on the residence of the presumptive clientele. It might be expected that planning for hospital care follows a similar pattern, that the "neighborhood hospital" should be the rule. The *Hospital Discharge Study* shows, however, certain unexpected facts: People do not necessarily go to the hospital located close to their residence; even if the neighborhood provides sufficient hospital facilities, they migrate on a city-wide scale. In New York City not only neighborhood but even borough boundaries are disregarded. Migration of this type is more common among the patients of voluntary hospitals than among those in municipal hospitals, but it exists in the latter group too. The fact that personal reasons and not the availability of proper local hospital accommodations influence a patient's choice of a hospital is best illustrated by the following example: Health area 59 in Manhattan has within its boundaries six general and one special hospital, some of them of a nation-wide reputation, with a total of 1,345 beds. The area has a population of 14,256 persons, 2,748 of whom needed hospital care during the year studied. These 2,748 persons went to 73 hospitals all over New York City; only 665 of them went to the seven institutions if the area.

Planning of health facilities for the population in the future will have to consider the hospital no longer as an institution per se, but administratively, as a link in a chain of health services, and, geographically, as but one part, though an important part, of the regional health service. If and when plans, which coordinate institutional bed care, outpatient care, health education and other weapons of sickness prevention, are realized, the somewhat chaotic flow of patients will cease and give place to organized care on regional principles. The first condition for a successful planning of the health service of the future is the knowledge of the strength of the enemy one will have to face, that is, the occurrence of diseases in a given population.

In summary, regularly reported data on the major causes of illness in a community are important:

(a) To shew the changing need for medical, nursing and hospital care, resulting from improved methods in preventing and curing diseases, and from the changing age composition of the population;

(b) To shew the needs of groups of the population characterized by special demographic and economic features; and

(c) To aid with the preparation and maintenance of a master plan of medical, nursing and social care for the sick in a community, a need that becomes more and more urgent since chronic conditions, with long-range and complex needs steadily gain in importance in the total disease picture of the population.

A CRITICAL APPRAISAL OF BUSINESS STATISTICS*

By JOHN S. PERKINS, *Assistant to the President*
Boston University

INDUSTRY has long followed Walter S. Gifford's dictum that "Business management is based upon facts." And the "facts" which the ordinary business executive is concerned with can be described as ninety per cent internal and ten per cent external. Teachers and writers on business statistics, on the other hand, have not yet come around to this understanding of the role which statistics play in the actual management of a business firm. In contrast with the business executive's point of view, the statistics and data that teachers and writers deal with can be described as ninety per cent external and ten per cent internal. Thus, it can be said that business statistics is not now fully meeting the business man's specialized needs in the statistical field.

For some time there has been a growing recognition in academic circles that the ordinary business statistics course does not bridge the gap between statistical technique and the actual use of statistics in business. Professor Theodore H. Brown of Harvard University warned against the disparity between the "preaching" and "practicing" of business statistics at a meeting of the American Statistical Association in December, 1936. On this occasion he stated that business statistics courses have been of little assistance in solving the specific problems that are actually faced by the average business man or by the young man in the position in which he will start his business career.[1]

More recently there have been other indications that this shortcoming is being recognized. Writing in the 1940 Summer Number of the *Harvard Business Review*, in a penetrating and sagacious article, "Statistics Takes a Second Breath," Professor Charles A. Bliss reported the result of his analysis of seventeen recent statistical textbooks, concerning which he stated, "It is very doubtful if such an array of outstanding (statistical) books, ever appeared in so short a period before." In his analysis of the contents of these books, Professor Bliss took pains to stress the significant fact that in only one book, Riggleman and Frisbie, *Business Statistics* (McGraw-Hill Co., New York, 1938) is material presented relating to actual statistical practices involved in the internal management of an individual business firm.

Another indication was noted at a discussion meeting of teachers of

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 27, 1941.

[1] This JOURNAL, Vol. 34, No. 206, June 1939, pp. 299 and 302.

statistics held during the Annual Meetings of the American Statistical Association in Philadelphia in December, 1939. Here the question was raised as to whether or not more attention should be given to internal statistics in college business statistics courses. According to the testimony of the representative group of teachers present at this meeting, the use and application of statistics in actual business management are seldom, if at all, incorporated in the average business statistics course. In reporting on the conference in the *Bulletin* of the American Statistical Association (May, 1939) Professor J. R. Stockton of Texas University summarized the opinion of those present when he said that "This represents an important omission in the offerings in strictly business statistics."

Here is a rich field awaiting cultivation. It is related to statistical and accounting techniques on the one hand, and to management theory and practice on the other hand. The area can most properly be called "Management Statistics."

It covers that field of statistical activity involving the analysis and use of internal operating data in specific departments and functions of individual business establishments. It is primarily the technique of collecting and analyzing management facts for the purpose of solving and controlling management problems.

Business statistics courses are not filling the need for training in this field. At its best, the ordinary business statistics course is an adaptation of the more simple and elementary statistical techniques to a pseudo-business setting based largely on general business and economic data. The average business statistics course now conducted provides no insight into the application of statistical and fact-finding techniques and principles in actual management practice. It bears little relationship to "statistics in business." College students in business administration go out with little or no appreciation of the nature, and use, and value of management facts in the actual internal administration of an enterprise. The statistical profession as a whole, statisticians, teachers, and writers, have virtually ignored the entire field.

The development of the field of Management Statistics will have to start from scratch. Teachers will have to prepare themselves in the field. Problem material must be developed. Literature in the field must be encouraged, and textbooks and course syllabi prepared. And preceding all this will have to come a framework, or classification, which can be used in sorting out and placing in logical position the various principles and practices as they are uncovered and explained.

The most important requisite is teaching leadership. To a large ex-

tent, teachers of business statistics have been unaware of Management Statistics simply because they have failed to keep in touch with the actual role played by statistics in business. They have been smoothing out curves of hypothetical business indices, while the business man has been trying to smooth out curves of machine loads, price lines, inventories, and the like. The result is that the student of so-called business statistics receives neither fish nor fowl. The statistical theory he gets is superficial, and the application of statistics to internal management is virtually non-existent.

The obvious answer is that business statistics teachers must re-establish contact with store and factory, with office and warehouse. Furthermore, this contact must be of a special nature. An occasional plant visit is not enough. They must go into the various departments of a business, see how the data originate, understand the units of measure and the technique of recording the information, and then follow them through various levels of authority and use and detail until they emerge in the form of a control report for top management. Because one is dealing here with the very nervous system of an individual business and with the confidential data which this implies, no mere speaking acquaintance with industry will suffice. There must be a real merging of interests and desires on the part of business management and teachers of statistics based on mutual understanding and desire to contribute something of value.

The business man himself should help if the field of Management Statistics is to be adequately developed. His most important contribution is to cooperate with the academic and professional world in supplying facilities for basic investigations of Management Statistics.

He can also cooperate by helping to determine the direction in which attention should be given to the field. Every executive worth his salt in business today must rely substantially on statistical records, reports, and analyses for knowledge of his own business and as a basis for his plans and programs. The business man knows best his needs for control data and how the facts are gathered and analyzed and presented. For these reasons, he can be of considerable value and help to the research worker in analyzing this management nervous system to find out what it consists of, how it operates, and how it can be most effectively used.

The development of literature in the field is an important need. Of the many different categories of Management Statistics, market statistics, and sales forecasting and, to a lesser extent, personnel statistics, are the only types covered in present-day business statistics courses and texts with any degree of completeness. In accounting literature there

exists much of value dealing with that important branch of Management Statistics related to budget and cost statistics and reports. However, until the many aspects of Management Statistics can be brought together in one literature, little progress can be made in formulating a body of theory and practice in this field.

Management Statistics should be recognized now as a definite branch of the statistical profession. This field represents one of the most important techniques in present-day business management.

The modern business man bases his decisions on *facts*. In one sense his staff organization exists for the sole purpose of supplying these facts in usable form. Without a system of organized statistics and reports, management would be paralyzed. There could be little delegation of authority without Management Statistics. In fact, the effectiveness of a firm is largely dependent on the choice and use of management statistics and analyses which keep the executive advised as to his course and his route. Yet little recognition has ever been given in professional and academic circles to this important field.

Full understanding and utilization of Management Statistics is especially necessary in a period of intense business activity and expanding operations. Under such conditions, personal oversight of the many and varied aspects of management is next to impossible. Decisions must be made quickly and accurately and it is absolutely essential that the executive have all the facts at his command, properly analyzed, interpreted, and presented.

From a vocational standpoint, training in the nature and scope of Management Statistics would pay almost immediate dividends. Most college graduates starting out in business must work to a considerable degree with records and reports dealing with various phases of a business. A person who has viewed the entire field of Management Statistics and is familiar with the nature and use of management records and reports is better qualified right from the start to display that added ability which marks him from the rest, thus laying the foundation for future executive success. One of the best ways of getting ahead in any organization is to display the ability to handle figures well, and to obtain and present effectively information on any type of management problem. After the basic traits of knowledge and judgment, this ability depends largely on a familiarity with Management Statistics.

Furthermore, attention to the field of Management Statistics would improve the use and value of management statistics themselves. There are instances in almost every company where reports and statistics are compiled in great detail by statistical departments without any one

finding any use for the figures thus prepared. Mr. Emil Hoefsos long ago summed up the situation by saying that, "Unfortunately, in too many *businesses, the executives are not statisticians, and* the statisticians are not business men."[2]

Business executives themselves have not even begun to utilize the full potentialities of Management Statistics as instruments of administration and control. Of course, management statistics and reports do not take the place of ability and judgment, but they are vital requirements for the effective use of ability and judgment. They are the methods of communication, the measures, and to a large extent, the means of interpreting happenings without which ability and judgment would be rendered impotent.

An automobile will not get one to a destination by itself. Only when the vehicle is used properly will it proceed to its goal. And the better the vehicle and the better trained the chauffeur to operate it, the more pleasant, safer, and surer will be the journey. By the same token, a sales analysis, or a production report, or a budget statement lying on an executive's desk, will not add one iota to the effective management of a firm. Only when the facts thus disclosed are used by the executive will an advantage be gained. And the advantage will be greater by the same measure that the report is effectively prepared and that the executive knows how to obtain the utmost value from the data presented in the report.

[2] "*Advertising and Selling,*" Vol. VIII, No. 5, pp. 21, December 20, 1926. (Also cited in Saunders and Anderson, *Business Reports*, 1st Ed., McGraw-Hill, New York, 1929.)

THE MYTH OF THE SECURITY AFFILIATE

By George W. Edwards
College of the City of New York

The decade of the thirties witnessed the legal divorce of investment from commercial banking. The Banking Act of 1933 forbade incorporated commercial banks from operating so-called investment affiliates and forced private banks to abandon either their deposit or their security operations. The Act sought to convert the function of American banking from an integrated to a specialized type.

This legislation rested on the belief that financial institutions combining both commercial and investment operations were against public interest. The hearings conducted by Congressional committees disclosed irregularities and abuses by the affiliates and by the private banks extending both investment and commercial credit. Subsequent financial literature, using the hearings as source material, has uniformly condemned integrated banking. In consequence this type is anathema to the general public. And yet the hearings did not face the fundamental issue of comparing the securities sponsored by integrated investment banks as against those floated by specialized investment banks. Save for a fragmentary table, furnished by a bank and not by the government, the hearings[1] did not even show the classes of securities floated by the various types of investment houses. The hearings presented no statistical analysis of the relative market results of the securities sold to the public by the non-affiliates as compared with those sponsored by the affiliates or by the private banks.

This article[2] undertakes a comparative analysis of the domestic corporate bonds floated by the various types of investment banking institutions. The analysis includes 2,633 bonds with a par value of 19,521 million dollars which provide a comprehensive base for deriving conclusions. These were all the bonds outstanding at the end of 1936 on which complete data could be obtained.[3] These bonds are studied in relation to the various types of sponsoring investment banks. They are first grouped according as they are (1) non-affiliates or those performing no commercial banking operations, (2) affiliates or those structurally related to incorporated commercial banks, and (3) private houses or

[1] *Hearings on the Operation of National and Federal Reserve Banking System*, Part 2, pp. 208-09. United States Banking and Currency Committee (Senate) 71st Congress, 3rd Session.

[2] The conclusions in this study do not apply to shares, foreign corporate or domestic or foreign government bonds floated by the investment banks.

[3] This analysis is part of a larger study begun early in 1938, and the end of 1936 was the latest date for which annual figures were available. The study omitted all bonds issued after the Banking Act of 1933 went into operation.

unincorporated houses engaged in both deposit and security banking. These banks are further classified by size, as large and small. The large banks include all the major investment houses in the period before the passage of the Banking Act of 1933, as listed by the *Wall Street Journal*[4] in its compilation of security flotations for the years from 1927 to 1932 inclusive. The small banks include all the remaining financial institutions.

RELATIVE IMPORTANCE OF INTEGRATED AND SPECIALIZED INVESTMENT BANKING

The greater part of the corporate bonds was sponsored by non-affiliates (Table I). They floated two-thirds of the amount of bonds, private banks a fifth, while the affiliates accounted for less than a seventh. The affiliate was therefore the least important institution in the new corporate bond market.

TABLE I
AMOUNT OF CORPORATE BONDS FLOATED BY INVESTMENT BANKS
(Per cent of total amount of bonds)

Investment bank	Large	Small	Total	Total
				Million dollars
Non-affiliate	50.01	15.67	65.68	$12,819.5
Affiliate	11.20	2.40	13.60	2,664.5
Private bank	19.96	0.76	20.72	4,037.3
Total	81.17	18.83	100.00	$19,521.3

The large houses dominated the corporate bond market. They accounted for over four-fifths of the total amount of bonds in the study. The large banks were far more important than the small banks in all three groups of houses. The large banks accounted for almost all the issues floated by the private banks, but were relatively less important in the case of the non-affiliates.

There were certain interesting variations from these total figures when the bonds were separated according to class of issuer (Table II). Bonds may be conveniently grouped according to issuer, as utility, rail, industrial, real estate and financial. In the utility fields the non-affiliates were especially important. They floated over three-quarters of the electric light, natural gas, water and utility holding company bonds. On the other hand, compared to their average for bonds as a whole, the affiliates had a high proportion in the weak traction and toll fields. The private banks were important only in the strong telephone field,

[4] *Wall Street Journal*, March 22, 1933, p. 10.

where they sponsored over half of the financing. The private banks also had more than their proportionate share of railroad financing. They offered over 70 per cent of the terminal bonds and over 28 per cent of the general railroad obligations. The equipment bonds, due probably to the

TABLE II

DISTRIBUTION OF CLASSES OF BONDS ACCORDING TO INVESTMENT BANKS

(Per cent of amount of bonds in each class)

Class of issuer	Non-affiliate		Affiliate		Private bank		Total	Total amount
	Large	Small	Large	Small	Large	Small	Per cent	Million dollars
Utility								
Electric light	00.5	15.5	10.8	1.7	0.8	1.7	100.0	$4,320.5
Natural gas	28.7	50.0	16.7	3.7	—	—	100.0	168.3
Telephone	30.1	5.7	7.2	1.0	55.0	—	100.0	800.0
Toll	40.0	20.1	13.3	11.0	2.7	—	100.0	63.8
Traction	26.5	7.0	27.8	6.1	28.8	4.2	100.0	055.0
Water	51.1	41.5	4.6	2.0	—	—	100.0	210.8
Holding	75.1	10.2	10.7	3.5	—	0.5	100.0	1,502.4
Total	60.0	14.2	11.8	2.5	14.1	1.4	100.0	$7,880.3
Railroad								
General	48.8	12.1	8.8	1.8	28.4	0.1	100.0	$7,909.0
Equipment	31.7	14.2	20.2	0.4	27.5	—	100.0	401.8
Terminal	11.8	15.0	0.0	0.3	71.4	—	100.0	327.0
Total	40.0	12.3	0.4	1.0	30.0	0.1	100.0	$8,757.5
Industrial								
Manufacture	52.0	15.4	24.0	3.3	3.4	0.4	100.0	$1,231.3
Extractive	30.2	21.5	18.8	4.1	10.1	0.3	100.0	344.1
Service	57.0	30.0	4.1	1.7	0.3	—	100.0	285.2
Trade	20.1	40.7	0.5	15.5	17.2	—	100.0	50.4
Total	40.7	20.4	10.8	3.0	0.2	0.3	100.0	$1,020.0
Real estate	21.5	02.2	7.2	3.5	0.0	—	100.0	$ 054.4
Financial								
Finance company	35.0	58.1	1.5	5.4	—	—	100.0	$ 33.1
Investment trust	65.4	15.5	4.7	—	1.6	12.8	100.0	177.0
Total	60.0	22.2	4.2	0.9	1.4	10.7	100.0	$ 210.1

fact that they were sold through competitive bidding, were more evenly distributed among the various types of investment houses. The non-affiliated banks financed about two-thirds of the general railroad bonds. Of the industrial bonds, the affiliates financed more than their general average, but the private banks less than their general average. The former financed almost a quarter of these bonds, and the latter only 6 per cent. The affiliates were important in the manufacturing field, while the private banks accounted for approximately a fifth of the extractive and the trade issues. The small non-affiliated banks completely

dominated the uncertain real estate, finance company and investment trust bonds.

There were also significant variations in the distribution of the several classes of bonds, according to the large and the small houses. In the utility field the large houses accounted for an unusually heavy proportion of the electric light, telephone, traction and holding company

TABLE III

CLASSES OF BONDS FLOATED BY INVESTMENT BANKS

(Per cent of total amount of bonds)

Class of issuer	Non-affiliate		Affiliate		Private bank	
	Large	Small	Large	Small	Large	Small
Utility						
Electric light	27.0	22.0	21.4	15.8	11.0	40.0
Natural gas	0.5	2.0	1.2	0.2	—	—
Telephone	2.8	1.8	3.0	1.8	13.1	—
Toll	0.3	0.6	0.4	1.4	0.1	—
Traction	1.7	1.0	8.3	8.0	4.0	18.0
Water	1.2	3.0	0.5	2.0	—	—
Holding	12.1	5.2	7.0	11.7	—	5.8
Railroad						
General	40.0	31.0	32.0	30.3	68.5	7.2
Equipment	1.5	2.2	5.0	0.3	3.3	—
Terminal	0.4	1.7	0.2	0.2	0.0	—
Industrial						
Manufacture	6.7	0.2	13.9	8.5	1.0	3.4
Extractive	1.3	2.4	2.9	3.0	1.7	0.0
Service	1.7	3.4	0.5	1.0	—	—
Trade	0.2	0.8	0.01	2.0	0.2	—
Real estate	1.4	13.4	2.1	11.8	0.1	—
Financial						
Finance company	0.1	0.6	—	0.4	—	—
Investment trust	1.1	0.0	0.4	—	0.1	15.4
Total—per cent	100.0	100.0	100.0	100.0	100.0	100.0
Total—amount (million dollars)	$6,715.7	$3,015.0	$2,187.9	$105.8	$3,800.0	$147.0

bonds. On the other hand, the small houses financed a large proportion of natural gas, toll, and water bonds. The large houses completely dominated the flotation of railroad bonds. Over four-fifths were sponsored by the major houses. On the other hand, the small houses gained more than their share of industrial financing, particularly service and trade companies. The small non-affiliates accounted for almost three-fifths of the real estate and finance company bonds. The minor private banks, while financing only a small fraction of the total amount of bonds, accounted for an eighth of all the investment trust bonds.

As a result of these different distributions, according to the class of issuer, there were variations in the composition of the total bonds

floated by the several types of investment banks (Table III). While electric light and railroad bonds naturally constituted the bulk for each type of bank, the percentage ranged widely. Almost three-fifths of the bonds floated by the large private banks were general railroad obligations. Of the bonds sold by the large non-affiliates, the electric light and general railroad bonds constituted over two-thirds. These bonds composed about half of the issues floated by the small non-affiliates and by the large affiliates, and less than half of the bonds financed by the small affiliates. Almost half of the bonds of the small private banks were electric light issues.

TABLE IV
INVESTMENT RESULTS OF BONDS ACCORDING TO INVESTMENT BANKS
(Per cent)

Tests of investment results	Non-affiliate		Affiliate		Private bank		Total				
	Large	Small	Large	Small	Large	Small	Large	Small	Non-affiliate	Affiliate	Private bank
Coverage*	1.03	1.50	1.08	1.41	1.75	1.02	1.60	1.49	1.58	1.58	1.73
Price stability	57	62	50	60	40	57	56	61	59	59	40
Default											
Number	21	40	27	35	10	20	20	30	30	30	11
Amount	17	34	21	25	0	8	10	32	21	22	0
Net return†	5.40	5.75	5.42	5.71	4.80	5.20	5.32	5.73	5.58	5.54	4.04
Median yield‡	5.16	8.83	5.24	7.48	4.17	5.87					

* Times fixed charges earned.
† Date of issue.
‡ December, 1936.

Each type of bank specialized in a minor field. The large non-affiliates were interested in financing utility holding companies. Manufacturing bonds represented fourteen per cent of the issues of the large affiliates, while real estate bonds constituted 13 per cent of the small non-affiliate issues and 12 per cent of the small affiliate securities. Telephone bonds accounted for 13 per cent of the offerings of the large private banks, while the small private banks had a conspicuous proportion of investment trust issues.

INVESTMENT RESULTS

The second part of this article compares the market action of the bonds floated by the various types of houses (Table IV). This market action is judged by studying the fixed charge coverage, price stability, default, net return and median yield. Coverage was the number of

times that net earnings were available to meet the fixed charges on the bond of a corporation. In this study a ten-year average was taken, and in the comparatively small number of cases where the earnings record was incomplete, as long a period as possible was used. Price stability was computed by taking the variation between the highest and the lowest price of each bond for the period from 1930 through 1936, and dividing the difference by the highest price. Default was interpreted as failure to pay fully and unconditionally the interest or principal of the bond. These tests were applied to each bond separately and so the results refer to the number of the bonds of each type of bank and not to amount, except in the case of default where both number and amount were used. The net return was computed from the issue price, maturity and coupon of each bond. This net return was used as a basis for judging the quality of the bond at the time of its offering to the investing public. In general a high net return on a bond indicated poor grade, while a low net return reflected good quality. The median yield was the yield to maturity of the middle group of bonds when ranged according to numbers and in the study the end of December, 1936, was taken as the basic date.

There was practically no difference in the investment results of the bonds floated by affiliate and non-affiliate banks. In fact there was a remarkable uniformity in the market experience of these two groups. They both showed the same coverage of 1.50 times, the same price range of 59 per cent and almost the identical percentage of default of 30 per cent for the total number and 21 and 22 per cent for the total amount. The bonds of the private banks were far superior in every respect. The average coverage was higher, they were more stable in price and the percentage of default was far smaller at only 11 per cent for the total number and 9 per cent for the total amount.

There was a significant difference between the bonds of the large and of the small houses. The bonds of the large banks were much better than the issues of the small banks. The coverage on the former was higher, their price was somewhat more stable and the percentage of default was much lower. The superiority of the bonds of the large houses is confirmed when they are studied separately according to each type of bank, for in almost every case the bonds of the large banks had better coverage, steadier price range and smaller percentage of default.

The above conclusions on coverage are confirmed by a study of the separate classes of bonds (Table V). This analysis is necessary, since coverage varies with the different classes of bonds. There was no pronounced difference in the coverage of the various classes of bonds

floated by the non-affiliate and the affiliate bonds. The traction, utility, holding and real estate bonds of the non-affiliates were somewhat better protected, while the electric light, railroad and manufacturing bonds of the affiliates had slightly better coverage. In every case the bonds sponsored by the private banks had decidedly stronger coverage. This was true of the electric light, telephone and manufacturing bonds, while only their traction bonds were weaker. In the greater number of cases for each class of bond, the large houses had better protection than the small concerns. Specifically 11 of the groups of bonds of the large houses had better coverage as against only 4 of the smaller houses.

TABLE V
COVERAGE OF CLASSES OF BONDS FLOATED BY INVESTMENT BANKS
(Time Fixed Charges Earned)

	Non-affiliate		Affiliate		Private bank	
	Large	Small	Large	Small	Large	Small
Electric light	2.03	1.06	2.20	2.09	2.73	1.85
Telephone	2.47	1.06	*	1.84	2.70	*
Traction	1.12	1.37	1.01	0.82	1.12	1.32
Utility holding	1.08	1.43	1.02	1.32	*	*
Railroad-general	1.41	1.29	1.84	1.40	1.50	*
Manufacturing	1.75	1.43	1.81	2.01	2.24	*
Real estate	1.31	1.06	0.85	1.09	*	*

* Insufficient number of bonds to evaluate computation.

A study of the net return and the median yield shows certain interesting features. There was practically no difference at the time of issue in the quality of the bonds floated by the non-affiliates and the affiliates. The net return on both these groups showed very little difference. There was a spread of only two points between the net return on the large non-affiliate bonds and that on the large affiliate bonds, only four points between the small non-affiliates and the small affiliates, and but four points between the total affiliates and the total non-affiliates. The quality of the bonds of the private banks was much higher than of those of both the non-affiliates and the affiliates. The bonds of the private banks sold on a yield of 60 points lower than those of the affiliates and 64 points below those of the non-affiliates.

The bonds of the large banks were of better grade than those of the small banks. This condition was true of all three types of banks. The large non-affiliate bonds were 35 points below the issues of the small non-affiliates; the bonds of the large affiliates were 29 points lower, and the large private banks, 31 points lower. The total bonds of the large

banks were 41 points below those of the total for the small banks. The bonds of the large houses brought more favorable results to the investing public than those of the small houses.[5] On the one hand, the issues of the large non-affiliates, the large affiliates and the major private banks, all showed net returns above the median yields, while, on the other hand, the bonds of the small non-affiliates, small affiliates and minor private banks all evidenced net returns below the median yields. In the case of the small non-affiliates and small affiliates, the return at date of issue was lower than the median yield indicating heavy losses to the investing public.

In conclusion this study shows that there is no statistical foundation for the belief that specialized banking was superior to integrated banking. There was practically no difference in the investment results of the bonds of non-affiliate and affiliate banks. In fact, the bonds of the private banks performing both commercial and investment operations were superior to those of the non-affiliate and affiliate banks. It would seem therefore that the Banking Act of 1933 separating commercial from investment banking did not rest on factual foundation. The belief in the utter financial depravity of integrated as compared with specialized investment banking is a myth and not a legend. A legend at least has an element of historic truth.

[5] This trend may be seen by studying the net return in relation to the median yield. This relationship shows in a general way whether the bonds in a group have appreciated or depreciated in value from their date of issue. A net return above the median yield indicates that the group of bonds have appreciated in value and that the investing public has gained from this group. Conversely a net return below the median yield means that the group has depreciated in value and that the investing public has lost on the group since the date of offering.

THE APPLICATION OF THE THEORY OF LINEAR HYPOTHESES TO THE COEFFICIENT OF ELASTICITY OF DEMAND

By M. A. Girshick
U. S. Department of Agriculture

Approximations to the standard error of the coefficient of elasticity of demand were given in this Journal by Henry Schultz[1] and Jacob L. Mosak.[2] In a more recent issue of this Journal, H. Gregg Lewis[3] has shown that under certain conditions, the distribution of this coefficient can be approximated by the normal curve.

The difficulty in applying the results obtained by the above authors to any given problem is that both the standard error formulae and the distribution involve population parameters whose values are seldom known. The substitution of sample estimates for these parameters is from a theoretical point of view a doubtful procedure, especially if the estimates are based on a small sample.

Here it is proposed to deal with the coefficient of elasticity of demand from the point of view of the theory of linear hypotheses. The statistical inferences drawn concerning this coefficient will not depend on any a priori knowledge of the values of population parameters.

In the following discussion it will be assumed that the mean or expected value of a normally distributed quantity variate X_1 is linearly related to $X_2, X_3, \cdots, X_k$, where X_2 is price and $X_3, \cdots, X_k$ are other variables. More specifically, it will be assumed that for given values of X_2 to X_k, X_1 is normally distributed with variance σ^2 and mean

$$\overline{X}_1 = \beta_1 + \beta_2 X_2 + \cdots + \beta_k X_k,$$

where $\beta_1, \beta_2, \cdots, \beta_k$ are the population regression coefficients.

Since for a demand function $\partial\overline{X}_1/\partial X_2$ is negative, the partial coefficient of elasticity of demand η will be defined as:

$$\eta = -\frac{\partial\overline{X}_1}{\partial X_2}\frac{X_2}{\overline{X}_1} = \frac{-\beta_2 X_2}{\beta_1 + \beta_2 X_2 + \cdots + \beta_k X_k}. \tag{1}$$

Suppose that a sample of N independent observations has been obtained on the quantity variate X_1 with corresponding observations on the variables X_2 to X_k. On the basis of the information supplied by this sample, the following statistical problems may be raised:

[1] "The Standard Error of the Coefficient of Elasticity of Demand," March 1933, pp. 64–69.

[2] "The Least Squares Standard Error of the Coefficient of Elasticity of Demand," June 1939, pp. 353–361.

[3] "On the Distribution of the Partial Elasticity Coefficient," September 1941, pp. 413–416.

Problem I: Estimate the elasticity of demand η for assigned values of X_2 to X_k.

Problem II: Test the hypothesis that $\eta=\eta_0$(η_0 being a definite number) when X_2 to X_k have some assigned values.

Problem III: For given values of $X_2, X_3, \cdots, X_k$ find fiducial limits for η.

Problem IV: For given values of $\eta, X_3, \cdots, X_k$, find fiducial limits for X_2.

Estimating the elasticity of demand. The problem of estimating the coefficient of elasticity of demand is usually answered by substituting for the β's in formula (1) the corresponding least squares regression coefficients, and for X_2 to X_k the specific values assigned to them. This method yields a maximum likelihood estimate of η. This follows from the fact that a maximum likelihood estimate of a function of population parameters is equal to the same function with the parameters replaced by their maximum likelihood estimates.

Testing the hypothesis that $\eta=\eta_0$. In order to answer problem II let it be assumed that the values assigned to $X_2, X_3, \cdots, X_k$ are $X_2', X_3', \cdots, X_k'$ respectively. The hypothesis to be tested is that

$$\frac{-\beta_2 X_2'}{\beta_1+\beta_2 X_2'+\cdots+\beta_k X_k'}=\eta_0,$$

or

$$\beta_1+\beta_2 Z_2+\cdots+\beta_k Z_k=0 \tag{2}$$

where

$$Z_2=X_2'\left(\frac{1+\eta_0}{\eta_0}\right) \quad \text{and} \quad Z_j=X_j' \qquad (j=3,\cdots,k).$$

The hypothesis as stated in (2) is an example of a general class of statistical hypotheses known as "linear" hypotheses. The first comprehensive discussion of such hypotheses was given in 1935 by the Polish statistician St Kolodziejczyk.[4] For a full discussion of the theory of testing linear hypotheses, the reader is referred to Kolodziejczyk's paper and also to a paper by Palmer O. Johnson and J. Neyman.[5] For the purpose of this paper it will suffice to give the details for testing the hypothesis expressed by equation (2).

Let $X_{i\alpha}$ stand for the αth observation of the ith variable ($i=1, 2, \cdots, k$; $\alpha=1, 2, \cdots, N$). Let $\bar{X}_i$ be the sample mean of the ith variable and $a_{ij}=\sum_{\alpha=1}^{N}(X_{i\alpha}-\bar{X})(X_{j\alpha}-\bar{X}_j)$. Furthermore, let

[4] "On an Important Class of Statistical Hypotheses," Biometrika, Vol. XXVII, 1935.

[5] "Tests of Certain Linear Hypotheses and Their Application to Some Education Problems," *Statistical Research Memoirs*, Vol. I (1936) Department of Statistics, University of London, University College.

$b_1, b_2, \cdots, b_k$ be the least squares estimates of $\beta_1, \beta_2, \cdots, \beta_k$ respectively. In terms of these quantities it can be shown that the statistic appropriate for testing hypothesis (2) is given by

$$Q = \frac{\overline{X}_1 + b_2(Z_2 - \overline{X}_2) + \cdots + b_k(Z_k - \overline{X}_k)}{s\left[1/N + \sum_{i=2}^{k}\sum_{j=2}^{k} C_{ij}(Z_i - \overline{X}_i)(Z_j - \overline{X}_j)\right]^{1/2}} \tag{3}$$

where s is the *standard error of estimate* [i.e.

$$s = \sqrt{\frac{a_{11} - b_2a_{12} - \cdots - b_ka_{1k}}{N - k}}$$

and C_{ij} is the cofactor of the element a_{ij} in the matrix (a_{ij}) divided by the determinant of the matrix (a_{ij}), $(i, j=2, 3 \cdots, k)$].

From the general theory of testing linear hypotheses, it follows that the quantity in (3) has a Student's t distribution with $N-k$ degrees of freedom.[9]

Therefore, to test the hypothesis that $\eta=\eta_0$, the following procedure is to be used. Calculate Q in (3). Find from published tables that value of t which is exceeded in random samples say five times in a hundred. (That is, find the value of t for the .05 level of significance with $n=N-k$.) If the calculated value of Q (disregarding sign) exceeds the value of t given in the table, the hypothesis that $\eta=\eta_0$ is rejected. If Q does not exceed this value of t, the hypothesis is not rejected.

Fiducial limits for η. The fiducial limits for the coefficient of elasticity of demand can be obtained from the following considerations: It will be noticed that the quantity Q in formula (3) has a Student's t distribution *no matter what* η (*contained in the definition of* Z_2) *is*. Consequently one can choose a level of significance say .05 and setting $n=N-k$, find that value of t (which may be designated by $t_{.05}$) which corresponds to the level of significance chosen. Then setting $Q^2=t_{.05}^2$ one can solve for η. Since the equation $Q^2=t_{.05}^2$ is of the second degree in η, the solution will yield two values of η. These two values are the required (upper and lower) fiducial limits for η. The interval defined by these two values is sometimes known as the "confidence" interval.

[9] The truth of the above statement can also be seen from the following considerations. Let $u = b_1 + b_2Z_2 + \cdots + b_kZ_k$. Then if the hypothesis is true, the expected value of u is zero. Now a straightforward calculation will show that σ_u^2, the variance of u, is precisely the square of the quantity given in the denominator of (3) with σ^2 substituted for s^2. Moreover, since u is a linear function of the least squares regression coefficients, it is distributed normally and independently of s^2 (since each b is distributed normally and independently of s^2). The quantity u^2/σ_u^2 has a chi-square distribution with 1 degree of freedom and $(N-k)s^2/\sigma^2$ has a chi-square distribution with $N-k$ degrees of freedom. Hence, the square root of the ratio of u^2/σ_u^2 to s^2/σ^2 (which is the expression given in (3)) has Student's t distribution with $N-k$ degrees of freedom.

The "confidence" interval thus obtained for the coefficient of elasticity of demand has the following meaning. If whenever the statistician is sampling a variate X_1 (which has the distribution defined on page 233) he calculates the fiducial limits for η and makes the statement that the true value of η lies in the interval thus defined, he will in the long run be right in 95 per cent of the cases. Or to put it in another way, if the statistician obtains a great many samples of a variate having the distribution attributed to X_1 and if he calculates the above intervals for each of the samples, then he may expect that 95 per cent of these intervals will cover the true value of η.

Fiducial limits for X_2. The statistician may in some cases be interested in estimating that price (X_2) which, for specified values of the remaining variables, corresponds to a given coefficient of elasticity of demand. This estimate may be obtained by substituting the b's for the corresponding β's in equation (1) and solving for X_2. The fiducial limits for X_2 for given values of η, $X_3, \cdots, X_k$, can be obtained in the same manner as outlined for obtaining the fiducial limits for η. That is, one sets $Q^2 = t_{.05}^2$ and solves for the two values of X_2 which this equation yields.

In dealing with the coefficient of elasticity of demand, the economist is often concerned with its value only at the point of averages of all or a subset of the independent variables. For this purpose formulae for the fiducial limits of the coefficient of elasticity and X_2 are derived below in an explicit form.

In equation (2) set $X_j' = \overline{X}_j (j = 2, 3, \cdots, k)$. Then Q takes on the simple expression

$$Q = \frac{\overline{X}_1 + b_2\overline{X}_2/\eta}{s[1/N + C_{22}\overline{X}_2^2/\eta^2]^{1/2}}. \tag{4}$$

Equating Q^2 to $t_{.05}^2$ and solving for η yields

$$\eta = \frac{b_2\overline{X}_1\overline{X}_2 \pm \overline{X}_2\sqrt{b_2^2\overline{X}_1^2 + (t_{.05}^2 s^2/N - \overline{X}_1^2)(b_2^2 - C_{22}t_{.05}^2 s^2)}}{t_{.05}^2 s^2/N - \overline{X}_1^2}. \tag{5}$$

The above two values of η give the required fiducial limits.

If $Z_j = \overline{X}_j$ for $j = 3, 4, \cdots, k$ and $X_2 = X_2'$, then Q reduces to

$$Q = \frac{\overline{X}_1 + b_2\left[X_2'\left(\frac{1+\eta}{\eta}\right) - \overline{X}_2\right]}{s\left[1/N + C_{22}\left\{X_2'\left(\frac{1+\eta}{\eta}\right) - \overline{X}_2\right\}^2\right]^{1/2}}. \tag{6}$$

Equating Q^2 to $t_{.05}^2$ and solving for η the result is

$$\eta=\frac{X_2'\left[(b_2^2-t_{.05}^2s^2C_{22})(X_2'-\overline{X}_2)+b_2\overline{X}_1\right]}{(t_{.05}^2s^2/N-\overline{X}_1^2)-\left[(b_2^2-C_{22}t_{.05}^2s^2)(X_2'-\overline{X}_2)^2+2b_2\overline{X}_1(X_2'-\overline{X}_2)\right]}$$

$$\pm\frac{\pm X_2'\sqrt{b_2^2\overline{X}_1^2+(t_{.05}^2s^2/N-\overline{X}_1^2)(b_2^2-C_{22}t_{.05}^2s^2)}}{(t_{.05}^2s^2/N-\overline{X}_1^2)-\left[(b_2^2-C_{22}t_{.05}^2s^2)(X_2'-\overline{X}_2)^2+2b_2\overline{X}_1(X_2'-\overline{X}_2)\right]}. \tag{7}$$

The above two values of η give the required fiducial limits.

If $\eta=\eta_0$ and $Z_j=\overline{X}_j$ for $j=3, 4, \cdots, k$, then the fiducial limits for X_2 are given by

$$X_2=\frac{\eta_0\left[(b_2^2-t_{.05}^2s^2C_{22})\overline{X}_2-b_2\overline{X}_1\right]}{(1+\eta_0)(b_2^2-t_{.05}^2s^2C_{22})}$$

$$\pm\frac{\eta_0\sqrt{b_2^2\overline{X}_1^2+(t_{.05}^2s^2/N-\overline{X}_1^2)(b_2^2-C_{22}t_{.05}^2s^2)}}{(1+\eta_0)(b_2^2-t_{.05}^2s^2C_{22})} \tag{8}$$

This is obtained by equating Q^2 in (6) to $t_{.05}^2$ and solving for X_2.

In case $k=2$, the quantity C_{22} in the formulae (4) to (8) has the value $1/a_{22}$ where, as defined above, $a_{22}=\sum_{\alpha=1}^{N}(X_{2\alpha}-\overline{X}_2)^2$.

Frequently, the variable X_2, instead of being the price of the quantity demanded, is a *known* function of the price. Thus, for example, X_2 may be given as $1/Y$ or $\sqrt{Y}$ where Y is the price of X_1. If it is assumed then that $X_2=f(Y)$, the coefficient of elasticity of demand is given by

$$\eta = -\frac{\partial\overline{X}_1}{\partial Y}\,\frac{Y}{\overline{X}_1} = -\beta_2\frac{dX_2}{dY}\,\frac{Y}{\overline{X}_1}$$

where β_2 is the population regression coefficient as defined on page 233.

In order to apply the results developed above to this situation it is only necessary to substitute in formula (3) for Z_2 the quantity

$\left(X_2'+\dfrac{dX_2}{dY}\,\dfrac{Y'}{\eta_0}\right)$ where $X_2'=f(Y')$ (Y' being a specified value of Y)

and dX_2/dY is evaluated at the point $Y=Y'$. The hypothesis that $\eta=\eta_0$ can then be tested and fiducial limits for η can be obtained in the same manner as outlined previously.

BUSINESS USES OF DATA BY CENSUS TRACTS AND BLOCKS*

By Vergil D. Reed, *Assistant Director*
Bureau of the Census

There has been entirely too much generalizing concerning the national market—and about the smaller markets of which it is composed. This may have been justified in the past through lack of dependable and comparable facts for these component parts of the total market but that is no longer a rational excuse. Today we can even take cities apart into their social and economic segments and see what makes them "tick." From the examination of a national market through a telescope we can turn to dissection and then examine the sections under the microscope. We begin to know something more than the generalities which have proven so grossly misleading when applied to specific cities or areas within them.

Sixty of our largest cities have been divided into census tracts. For 191 cities, including the tracted ones, housing data will even be published by blocks. This is in addition to all the facts published by cities, metropolitan districts, and minor civil divisions such as townships, parishes, beats, and towns.

Many data on population and housing are being published by tracts. Even the major items for retail trade can be made available by combinations of tracts. As examples of what can be done in these intra-city analyses of retail trade attention is called to four such studies made in the Census Bureau in connection with the Census of Business, 1935:

1. Geographic Distribution of Retail Trade in Chicago, Illinois,
2. Intra-City Business Census Statistics for Philadelphia, Pennsylvania,
3. Geographic Distribution of Retail Trade in Buffalo, New York,
4. Changes in Retail Trade in Buffalo, 1929, 1933, and 1935.

In the series of bulletins entitled "Statistics for Census Tracts," the following standard tables for housing and population will appear:

1. Population by race and nativity, and occupied dwelling units, by census tracts: 1940,
2. Age, race, and sex, by census tracts: 1940,

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 29, 1941.

3. Years of school completed, employment status, class of worker, major occupation group, country of birth, and citizenship, by sex, by census tracts: 1940,
3a. Nonwhite population by years of school completed, employment status, class of worker, major occupation group, and sex, by census tracts: 1940,
4. Dwelling units by occupancy status and race of occupants, by census tracts: 1940,
5. Value of owner-occupied dwelling units, contract or estimated rent of all dwelling units, and gross rent of tenant-occupied units, by census tracts: 1940,
6. Type of structure, state of repair and plumbing equipment, size of household, persons per room, radio, refrigeration and heating equipment, by census tracts: 1940,
6a. Dwelling units occupied by nonwhite households, by value or rent, size of household, persons per room, etc., by census tracts: 1940.

A table has been published in the First Series Population Bulletins showing the 1940 population of each tract (1930 population figures are shown where available). A map of the city by census tracts is included.

Additional information on the characteristics of housing by census tracts is being presented in a Supplement to the First Series Housing Bulletins.

For each of the 191 cities which had a population of 50,000 or over in 1930 a supplement to the First Series Housing Bulletins will be issued containing three tables:

1. Characteristics of Housing for the City: 1940,
2. Characteristics of Housing by *Census Tracts*: 1940,
3. Characteristics of Housing for Census Tracts *by Blocks*: 1940.

For additional guidance in the application of census data to the problems of a community it will be well to secure a copy of "Key to the Published and Tabulated Data for Small Areas." Many useful facts are tabulated which the Bureau does not feel justified in publishing because they are of interest only to a few groups or individuals. These tabulated data are available for the nominal charge of copying or reproducing. Other facts on housing, population, and retail trade are punched on cards but not tabulated. Special tabulations can be made at the expense of those who want them for tracts or combinations of tracts, wards, blocks, and enumeration districts, except that in the case

of retail trade it is not considered practicable to go beyond combinations of census tracts due largely to danger of disclosing individual operations. Far too few businessmen know the profit possibilities of this unpublished and untabulated information, available at a negligible fraction of the cost of collection. Enquiries regarding unpublished data are welcomed.

One of the greatest values of census tract and block data is that they can be used as a background to which facts collected by others can be related. The businessman can, in seeking a solution for his local problems, even collect additional information by these same areas to supplement the basic information furnished by the Census.

To the manufacturer, his advertising agency, and his marketing consultants the potential dividends from this small area material are particularly great. The detailed study of individual cities can be equally valuable as market analysis and as market synthesis. From a detailed knowledge of the treated cities he can construct accurately the characteristics of his total market. By conducting trial campaigns in carefully chosen trial markets the laboratory method is effectively applied to marketing. These trial markets no longer need be chosen by rule of thumb only to prove nonrepresentative and disappointing. They can be chosen on the basis of detailed analysis and the conditions may be controlled as desired.

Most market surveys must be on a sample basis—and on a "thin" sample at that. In such sampling tract and block data assure us of adequate controls, proper stratification, representativeness of the universe being sampled, and either consistency or the known causes of inconsistency. These are guides long overdue.

Where do families live? How large are they in various parts of the city? What kind of homes do they live in? Do they own or rent them? Where are the declining areas, the growing areas? What will be the direction of development? Where do the various racial and language groups live? Through what media, type of appeal, and type of salesman can they be reached? What is the purchasing power in the various areas? What parts of the city are worth intensive sales effort—and which can best be left to competitors? Which retail outlets should carry your high quality and your low quality lines? Which should carry the entire line and which should carry only certain items? How much in quantity and quality can you reasonably expect to sell in a given city—or parts of it?

Knowing the answers to those questions—and they are there for most products—improves your salesmen's percentage of sales to calls and

increases the size of sales. *Sales quotas are reasonable expectancies of sales based on facts.* They are given in detail for large city markets. They can be made the basis for real selective selling. Certain products will—or will not—sell best in homes having mechanical refrigerators, electric lights, bathtubs, oil heat or radios. It is easy to find where such homes are concentrated and how many there are of them. Knowing keeps sales costs down.

Some goods have highly specialized markets. Others have universal coverage. Types of clothing worn are related to age, sex, and race composition of the population. Reading matter is directly related to educational level. With the new census questions on education, income, and work status, and more complete tabulations on age, sex, color composition, marital status, labor force data, and the entirely new housing census, marketing research has come of age.

Banks and real estate interests have long been users of census tract data and they will undoubtedly find the block tabulations equally valuable. The housing census particularly affords them a wealth of information which they have never had before and it breaks these data down into the small packages necessary for most effective use.

Types of dwellings and the relative proportion of different types are important indicators of loan values. Single-family houses surrounded by multi-family homes have a loan value different from a single-family house in a single-family neighborhood. Value of dwellings is partly determined by location in relation to stores, factories, gasoline stations, and other special types of properties. A decrease in families or population count in an area indicates a declining trend in property valuation. The existence or non-existence of adequate heating, cooking, lighting, and bath room facilities indicates the commercial possibilities of properties.

As a guide to intelligent appraisal of land values and building improvements, the population and housing statistics by small areas should be a godsend. Mortgage conditions will be shown by tracts including the number of homes mortgaged, outstanding indebtedness, holder of mortgage, and total interest rate.

One banker states that by capitalizing the rental values in any area he gets a very definite value for real estate in that area. He also says that he uses census facts for arriving at banking policies, for the banker must know where, as well as when, to expand or contract loans.

Public utilities are constant users of intra-city area statistics. Upon them they must formulate their programs and plans as to the extension of telephone, electric light, gas, and transportation service. A telephone

executive states that his company must build a central office building and fill it full of equipment with an investment of from two to four million dollars. That building must have a satisfactory location. The population and housing characteristics of a city determine the location. These concerns must know where vacancies are high, where they are low and in what types of dwellings. They must know how many families live in one-family dwellings, how many live in apartments, and how many live in other types of multi-family units. Some of the public utilities have done particularly fine jobs. Others are expecting to do similar jobs in making surveys of their areas based on census tracts. One example of these is the Brooklyn Market Survey, with its many tract maps, made by the Brooklyn Edison Company in 1936. With the new facts available for 1940, both the number and the usefulness of these surveys will undoubtedly increase.

Many of the telephone companies in the past have made special arrangements with the Bureau to secure tabulations by enumeration districts since it was found necessary to have data by areas smaller than census tracts in order to analyze specific telephone service areas.

A telephone engineer writes, "The announced plans of the U. S. Census Bureau to publish their field data by blocks for the larger cities were received enthusiastically by telephone engineers since such information will prove of tremendous value to this industry."

Wide-awake retailers long ago learned the value of proper location for their stores. Now any retailer desiring to determine the best location for different types of stores has available to him practically all the facts he needs from the census, except traffic, upon which to make his choice of a site. He can certainly know his potential customers better and find better ways to reach the types of consumer which he wishes to cover.

The department store and the chain store, through tract and block data, can more intelligently locate warehouses and branches, plan their delivery services, determine effective advertising appeals, and choose the media best adapted to reaching their customers.

Chambers of commerce have always had a deep interest in facts which reflected conditions and characteristics in various parts of their cities. This knowledge is fundamental to any understanding of real conditions and to improving them. From community betterment and health campaigns to campaigns for getting new factories, these intra-city breakdowns are serving these organizations.

One chamber reports using them to help young physicians choose a location having two specific characteristics: First, it had to be a growing neighborhood, and, second, it had to be a neighborhood offering

opportunities to become "company doctor" for one or more factories. This same organization has also used the tract data to help men choose locations for motion picture theatres, shoe stores, hardware stores, and an aviation plant.

In selecting and soliciting business or manufacturing concerns most apt to benefit the community and most apt to succeed in the locality census facts are always more convincing than unsupported glibness. There is no better approach than presenting incontrovertible figures on the labor force, living conditions, purchasing power, and property values.

An automobile club executive says, "It would take a volume to report to you all the uses which we make of your Census Tracts and Blocks. However, briefly, our entire Membership Department is based on census tract information. Our various territories for representatives to work are set up in accordance with census tracts and all prospects for membership are so classified. We use this information in evaluating our districts from an economic standpoint and it has been of great help in enabling us to make various research studies which we use in connection with the solicitation of membership. We are now using it extensively in a study we are making of the results obtained from automobile driver education which is being carried on in the local High Schools."

The three outstanding news media—radio stations, magazines, and newspapers—are constantly coming to us to learn more about the applications of our statistics to intra-city areas.

Radio stations in reporting to the Federal Communications Commission must sometimes furnish population or number of families within certain intensity bands. By breaking the tracts which fall on band limits into blocks, an excellent analysis can be made for this purpose. Those in Cleveland are certainly aware of these possibilities and have long used them through applying the results of their Real Property Inventory. Census facts can be used in any of the tracted cities for this purpose.

Magazines and newspapers in any analysis of their readers, in setting up circulation plans, and in assuring maximum returns from the advertising carried by them will find the tract and block data indispensable.

If there is any slogan which the newspaper publisher should take to his heart it is "Know thy city well to serve it best." Editorials, news articles, and promotion plans are dependent upon factual knowledge of the city and its parts.

The progressive newspaper publisher must be capable of feeding facts as well as life into his news, advertising, and circulation departments no

less than into his editorials. Every publisher is interested in one certain community (or group of such communities)—a city and its hinterland. The city is the yolk and the hinterland is the white of the egg from which newspaper profits hatch. That egg must, however, be a sound one fertilized with creative imagination and progressive management lest it rot rather than hatch. Trial and error methods mean low hatching scores even with the best of incubators. The census provides the means for grading and candling that egg and determining in advance the conditions which will assure the hatching and maturing of the chick.

A newspaper which has large-scale home deliveries in a tracted city can allocate its subscribers by these tracts, then study intensively their social and economic characteristics—their economic status, education, color-nativity, age, sex, occupational and industrial composition, and the nature of their homes and the equipment in them.

Within the bounds of this paper it has been impossible to treat in detail the methods of applying tract and block data to the solution of many specific business problems. Certainly the applications covered can be adopted or adapted with profit by any businessman who wishes to do a better job of evaluating the opportunities and problems existing in larger cities. Putting to work the basic facts which are made so easily available to him in packages conforming to his practical needs is now only a matter of choice. However, even census facts possess value only through use.

GENERAL PRINCIPLES OF TRACT DELIMITATION*

By C. E. Batschelet, *Geographer*
Bureau of the Census

NOW THAT THE ACTUAL enumeration of the Census of 1940 is completed and the results are being published, it seems a good time to consider the entire tract program and to lay down certain general principles which our experience has shown us should be followed in delimiting these areas.

Prior to the last census, the cost of all tract tabulations was defrayed by the tract sponsoring groups, and the Bureau of the Census merely constituted a "Board of Review" to determine whether the tracts were established in conformity with certain fundamental principles. However, as the Bureau now tabulates and publishes the tract statistics, our part in the program should no longer be merely a review of the tract layout, but we should take an active part in the delimiting of the areas to the end that uniformity may be achieved—not uniformity in the size or composition of the tracts, but uniformity in the numbering, boundary descriptions, and problems of a similar nature. For example, at the present time, there are probably seven or eight different numbering systems in use, which is not good practice.

In the light of these facts, the Bureau of the Census has outlined the following course of action which each city interested in a tract setup will be asked to follow:

(1) The submission to the Bureau of the Census of a cadastral map, preferably on a scale of 1,000 feet to the inch, showing the proposed tracts, as well as the blocks, street layout, and street names. This map should show the existing streets as distinct from proposed street plans.

(2) The submission to the Bureau of the Census of a list of the agencies, public and private, approving the tracts, as well as the names of any agencies disapproving them.

(3) After the maps are reviewed by the Bureau of the Census, they will be returned to the tract sponsor, and the necessary changes must be made to adjust the tract lines to the mutual satisfaction of the participating agencies and this Bureau.

(4) The submission to the Bureau of the Census of a statement from the Mayor, or other head of the city government, officially approving the tract plan.

(5) The publication of a tract outline map which can be used by

* A paper presented at the 103rd Annual Meeting of the American Statistical Association. New York, December 29, 1941.

local agencies for plotting statistical data, and the submission of a copy of this map to the Bureau of the Census.

(6) The publication of a tract street index and the submission of two copies of this index to the Bureau of the Census.

As soon as all of the above conditions have been met, the tract sponsor will be notified that the Bureau of the Census recognizes that the city is tracted.

At the Census of 1940 there were 61 tracted cities of which 55 were municipalities of 100,000 or more population. This leaves 38 cities of the 100,000 class still untracted. It should be the goal to have tracts established in all of these cities in time for the 1950 Census. As a guide in the delimiting of tract areas, the Bureau of the Census is glad to send to any city considering the establishing of tracts a copy of a map showing the enumeration districts used in the city at the Census of 1940 with the population reported for these enumeration districts.

Further there were 24 cities which had discovered that tracts are just as essential in the territory surrounding the city as in the central city itself and had established tract units in the extra-municipal area. The Bureau of the Census feels that tracts should be established in the metropolitan areas of all of the 100,000 cities, and has accordingly drawn up the following regulations governing the delimiting of tracts:

(1) Tracts in the areas surrounding cities should be limited to the metropolitan district as established by the Bureau of the Census or to the county in which the central city is located.

(2) The tracts should conform to the boundaries of the county political subdivisions, i.e., each minor civil division should form a complete tract or be subdivided into two or more tracts. It is realized that for some cities this policy would occasion changes in the tract boundaries from decade to decade because of changes in the minor civil division lines. However, this now occurs in the periphery of a city, where the tract boundaries change on account of annexations to or detachments from the central city.

(3) The delimiting of tracts in the metropolitan district should be under the direction of the tract sponsors in the central city working in close collaboration with county and city officials.

(4) The tract sponsor should submit to the Bureau of the Census a cadastral map showing the tracts by the subdivisions specified in paragraph 2. For use in this work, the Bureau of the Census will be glad to supply, upon request, copies of the county base maps on file.

A METHOD OF ANALYZING THE ELEMENTS OF FORECLOSURE RISK*

By Mortimer Kaplan
Federal Housing Administration

THE FUNCTION of practical research in mortgage finance is to provide a basis for lending policy which will circumscribe the three principal problems of risk bearing, namely, the incidence and magnitude of foreclosure expectancy, the incidence and magnitude of probable losses, and the standards of risk selection. The need for research and its scope was recognized by the officers of the Federal Housing Administration simultaneously with the inception of mortgage insurance operations and the program was provided for in its division of Research and Statistics. The justification and value of such studies to private lenders as well as to the Federal Housing Administration should be obvious in view of the widespread interest in FHA mortgages manifest in the growth of the insured mortgage portfolios of financial institutions.

Foreclosure risk is the first of the two components of mortgage risk which must be studied to determine how risk factors are associated with bad mortgage loans. In establishing the limits of foreclosure risk for salient elements of the mortgage transaction, the maximum limits of loss risk, the second component of mortgage risk, are also delineated in terms of the risk elements. Loss risk, of course, will depend wholly on the financial experience in the disposition of the foreclosed mortgage security.

The purpose of this paper is to present a method of analyzing the elements of foreclosure risk and of measuring the variations in risk in terms of a foreclosure risk index and a foreclosure expectancy rate. This method was applied in the forthcoming monograph on the foreclosure experience with insured mortgages based on the first five years of operation of the Mutual Mortgage Insurance program.[1] The monograph covers the first of the three problems of risk bearing in both its phases, namely, over-all foreclosure risk and the components of foreclosure risk.[2] In the interests of economy of space, only the major features of this technique will be treated.

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 27, 1942.

[1] "Foreclosure Experience with Insured Mortgages: A Report on the First Five Years of Operation of the Mutual Mortgage Insurance Program," Division of Research and Statistics, Federal Housing Administration, Washington, D. C.

[2] For a complete statement of the scope of this study consult "An Analysis of Foreclosure Risk," *Insured Mortgage Portfolio*, Federal Housing Administration, VI, 1, Third Quarter, 1941, p. 16.

In presenting the method of measuring the foreclosure propensity of the various elements of insured mortgage risk, a general statement describing the technique of analysis and its rationale will precede the more detailed description of the adjustments to this method which were necessary to utilize all of the foreclosed mortgages which had occurred on mortgages insured in these first five years.

The implementation of the method of determining the variations in foreclosure risk associated with the separate risk characteristics reflected in the mortgage transaction, such as the type of borrower, type of mortgage, and type of mortgage security, involves compiling samples of good and bad mortgages for the various risk elements which are selected for study, and the preparation of parallel frequency distributions for convenient class intervals.

These percentage distributions are then compared to reveal differences by the use of a per cent relative which is derived by dividing the percentage of foreclosed cases in each of the class intervals by the percentage of good mortgages in the same class interval.

The resulting relative, which may be called a foreclosure risk index, measures the risk of foreclosure for each of the groups in terms of the average foreclosure experience of all groups. Thus the values of the foreclosure risk index for any one of the classes which are less than 100—the average foreclosure risk for mortgages in all groups—indicate better-than-average foreclosure risks and the groups with index values greater than 100 are worse-than-average foreclosure risks.

The rationale of this method, briefly, is first to determine the presence of differences from assumed uniformity in a dichotomous classification of mutually exclusive categories of good and bad mortgages; and, second, to measure these differences in terms of an average foreclosure experience with the mortgages under observation. The degree to which these differences are significant in terms of a reliable sample can be determined by the application of accepted statistical tests of significance. This is essentially the basis of the statistical technique employed in the above-mentioned study.

Because of the limitations of the available data on good and bad insured mortgages, the method actually employed in the analysis of foreclosure risk is an approximation of the method described above. These limitations arise principally out of the fact that the typical insured mortgage is a long-term amortized loan, the first of which was made under the Mutual Mortgage Insurance Program in 1935, and that a relatively insignificant number of foreclosures has since occurred. This paucity of experience is responsible, in the first instance, for the

modification of the definitions of terms used and, in the second instance, for the necessity of the adjustment in the basic data.

From a contractual point of view a good mortgage is one which is paid off according to the terms of the mortgage instrument. Such a definition is obviously not convenient to use in an analysis of the kind of long-term amortized mortgages insured by FHA. It will be many years hence before a sufficiently adequate sample of this highly desirable classification of mortgages will be amenable to analysis.

On the other hand, from a financial point of view, a good mortgage might include one on which the gross income and proceeds realized on sale in case of foreclosure are sufficient to cover all losses including expenses. However, since risk of foreclosure and not risk of loss is the objective of analysis, this profit and loss definition is *ipso facto* inadequate.

In lieu of these two currently inadequate definitions of a good mortgage, a definition, which provides a reasonably good approximation of the conditions implicit in the first of the proposed definitions and which affords at the same time an adequate workable sample of cases, is to be found in the mortgages-in-force, that is, the mortgages in good standing as of a particular date. The adequacy of this definition rests on the assumption that the bulk of the mortgages made will survive to maturity.

This difficulty in formulating a useful definition of good mortgages is not encountered in the definition of bad mortgages. For purposes of an analysis of foreclosure risk, a bad mortgage can be taken as being synonymous with a foreclosed mortgage. Foreclosure is an articulate legal status and its legal definition is useful to the kind of empirical approach undertaken here.

Ideal basic statistics for a foreclosure analysis would be provided if all mortgages were made at approximately the same time, say within the same year, and if the mortgages-in-force and the foreclosures were derived from this original portfolio. The sampling problem would consist in selecting all or a representative sample of those mortgages which survived during this period and all or a representative sample of those mortgages which were foreclosed. Under such conditions all mortgages in this ideal portfolio would be exposed to the same general risk factors.

The problem of selecting the samples of good and bad insured mortgages was automatically determined by the nature of the universe of each of these categories. The surviving mortgages as well as the foreclosed ones come from mortgages made in each of the five years between 1935 and 1939. With respect to the mortgages-in-force, the ho-

mogeneity of the mortgages insured was affected principally by the circumstances under which the Mutual Mortgage Insurance program got under way and by the subsequent amendments to the National Housing Act.

At the inception of insuring operations, the applications for FHA-insured mortgages came principally from home owners interested in refinancing their mortgages under more favorable terms. A much smaller, almost negligible, proportion came from purchasers of new homes. This was to be expected since home building activity had not yet reacted to the stimulus of federal recovery legislation. As time went on, this situation was reversed and purchasers of new houses, i.e., homes less than a year old at the time of commitment of insurance, represented a preponderant proportion of mortgagors under the Mutual Mortgage Insurance program. In Table I these changes can be seen in bold relief.

TABLE I

PERCENTAGE DISTRIBUTION OF MORTGAGES INSURED ON TOTAL, NEW, AND EXISTING 1- TO 4-FAMILY HOMES, 1935-1939

Year Insured	1935	1936	1937	1938	1939
All homes	100.00	100.00	100.00	100.00	100.00
New homes	28.13	35.84	49.84	60.95	71.98
Existing homes	71.87	64.16	50.16	39.05	28.02

Up to the end of 1937 the gradual changes in the character of the mortgage security can probably be explained by the general recovery in economic conditions and particularly in building construction. The sharp reversal in the type of mortgage security after 1937 can be explained in large part by the amendments of February 3, 1938, to the National Housing Act, which gave special consideration to the mortgage financing of new homes by extending the maximum term to 25 years and by raising the maximum loan-value ratio to 90 per cent. Not only did the proportions as between new and existing homes change, as indicated in the previous table, but the volume of mortgages increased markedly during this period. Moreover, significant changes were reflected in the average characteristics of the mortgages made in each of the years of this period, such as the lower average valuations, the longer average maturities, and higher loan-value ratios. Table II presents the proportion of mortgages insured in each of the five years, as well as the distribution of foreclosures for the same period.

These factors not only affected the homogeneity of the mortgages-in-force but also affected the distributions of the foreclosures out of

which they had to come. The character of the distributions of the so-called good mortgages, or mortgages-in-force, would be largely determined by the mortgages made in the later years, while the character of the foreclosures would be largely determined by those mortgages made in the earlier years.

TABLE II

PERCENTAGE DISTRIBUTION OF MORTGAGES INSURED AND FORECLOSED ON TOTAL, NEW, AND EXISTING HOMES SECURING MORTGAGES INSURED, 1935–1939, BY YEAR INSURED

Based on mortgages foreclosed through June 30, 1940

Year insured	All homes Insured	All homes Foreclosed	New homes Insured	New homes Foreclosed	Existing homes Insured	Existing homes Foreclosed
1935	5.03	12.17	2.63	5.00	8.18	17.00
1936	10.58	31.07	10.01	25.57	24.00	35.57
1937	21.02	30.78	10.60	40.54	24.80	33.70
1938	23.40	17.22	25.24	24.04	21.22	11.15
1939	33.01	2.70	42.03	4.10	21.62	1.62
Total	100.00	100.00	100.00	100.00	100.00	100.00

In order to have an adequate sample, all foreclosures on mortgages insured during this five-year period, 1935–1939, had to be used. However, there are differences from the assumed uniformity of the distributions of good and bad mortgages which could be attributed to the factors responsible for the heterogeneity of the mortgages and the varying exposure periods prior to foreclosure, and not necessarily to differences which could be attributed to foreclosure risk. This heterogeneity characteristic of the good mortgages can be corrected by introducing a parallel heterogeneity in the sample of foreclosures—a somewhat unorthodox statistical concept whose meaning will become clear presently. This is achieved by adjusting the varying exposure periods prior to foreclosure to periods of comparable length.

The adjustment developed to compensate for these two factors is based on the monthly foreclosure mortality experience table for insured mortgages. This foreclosure mortality table is constructed on the basis of the same principles as mortality tables on human life used by life insurance companies in the calculation of an appropriate premium rate. Just as a mortality table is designed to answer the question on the number of insured lives at various ages who are expected to die within a given period, so the mortality table for foreclosures presents the number expected to be foreclosed of the number of mortgages insured for varying periods of time.

Since the mortgages were insured in varying amounts in separate

years, the *desideratum* is to reproduce a situation in which all the mortgages were insured at the same time, or roughly in the same year, and examine the character of the experience which could be expected at the end of the period. This can be achieved through the application of the foreclosure mortality tables.

It should be clear that a greater number of foreclosures would have resulted if all mortgages were made in the first year of the five-year

TABLE III

RANGE AND AVERAGE PERIOD OF EXPOSURE PRIOR TO FORECLOSURE AND CORRESPONDING FORECLOSURE RATES ON TOTAL, NEW, AND EXISTING HOMES INSURED FROM 1935 THROUGH 1939, AS OF JUNE 30, 1940

	Year Insured				
	1935	1936	1937	1938	1939
	Range in Months Exposed Prior to Foreclosure				
All cases	1-66 mo.	1-54 mo.	1-42 mo.	1-30 mo.	1-18 mo.
	Average Exposure Period Prior to Foreclosure				
All homes	41.34 mo.	33.27 mo.	25.50 mo.	18.05 mo.	12.13 mo.
New homes	41.24	31.01	24.30	16.84	12.10
Existing homes	41.36	33.90	26.00	18.90	12.18
Average exposure period prior to foreclosure for all homes					29.11 mo.
Average exposure period prior to foreclosure for new homes					22.37 mo.
Average exposure period prior to foreclosure for existing homes					28.76 mo.
	Foreclosure Mortality Rate for Average Exposure Period				
All homes	1.225%	.899%	.650%	.293%	.099%
New homes	1.130	.764	.591	.254	.128
Existing homes	1.257	.923	.651	.272	.091
Foreclosure rate for average exposure period for all homes					.514%
Foreclosure rate for average exposure period for new homes					.411%
Foreclosure rate for average exposure period for existing homes					.627%

period. The longer the period of exposure to risk, the greater is the proportion of foreclosures out of a given number of mortgages made at the same time. The rationale of the adjustment method employed is as follows: On June 30, 1940, the mortgages which had been insured during the calendar year 1939 could have been foreclosed in any month from January 1939 through June 1940. The range of the period of exposure to foreclosure is thus from 1 to 18 months. The average period of exposure prior to foreclosure for the foreclosures on mortgages made in 1939 is 12.13 months. The proportion of mortgages which could be expected to be foreclosed after an exposure period of 12.13 months is .099 of 1 per cent. This foreclosure rate is derived from the foreclosure mortality table.

Table III presents the range in months, the average period of exposure, and the corresponding foreclosure rates for mortgages on total, new, and existing homes made in the separate years between 1935 and 1939.

It is apparent that if the mortgages made in 1939 had been made in 1935, there would have been many more foreclosures according to the foreclosure mortality table. Since the foreclosure rate for the 1939 mortgages is, according to the preceding table, .099 of 1 per cent, and the rate for the 1935 mortgages is 1.225 per cent, then 12.374 times (1.225 ÷ .099) as many foreclosures could have been expected by June 30, 1940, if the 1939 mortgages were made in 1935. The rate of foreclosure for the 1935 mortgages is thus 12.374 times the rate for the 1939 mortgages.

TABLE IV

ACTUAL AND ADJUSTED FORECLOSURES ON MORTGAGES SECURED BY ALL, NEW, AND EXISTING 1- TO 4-FAMILY HOMES INSURED, 1935–1939

Based on Mortgages Foreclosed Through June 30, 1940

Year Insured		All homes*	New homes	Existing homes
1935	Actual foreclosures	266	40	211
1936	Actual foreclosures	664	246	418
	Expectancy adjustment factor	1.395	1.453	1.362
	Adjusted foreclosures	926	357	569
1937	Actual foreclosures	780	300	306
	Expectancy adjustment factor	2.151	2.030	2.200
	Adjusted foreclosures	1,691	702	890
1938	Actual foreclosures	368	297	131
	Expectancy adjustment factor	4.533	4.481	4.021
	Adjusted foreclosures	1,668	1,063	905
1939	Actual foreclosures	59	40	10
	Expectancy adjustment factor	11.034	8.898	15.510
	Adjusted foreclosures	651	360	205
	Total actual foreclosures	2,137	902	1,175
	Total adjusted foreclosures	6,106	2,617	2,579

* The discrepancies between the values of the expectancy factors for all homes in this table and those in the text are explained by the somewhat different definitions of new and existing homes used in the construction of the mortgage mortality table. The differences in rates are, however, insignificant.

The value, 12.374 is called an "expectancy adjustment factor." This is used to convert the actual foreclosures on mortgages made in 1939 to what could be expected to prevail if these mortgages were made in the first year of the five-year period.

Expectancy adjustment factors are computed for the foreclosures on the mortgages insured in each year between 1936 and 1939. For the foreclosures on mortgages made in 1935, they are unity because it is this year which is used for the base foreclosure experience.

Actually, of the mortgages made in 1939, only 59 mortgages on 1- to 4-family homes were foreclosed by the end of June 1940. On the basis of the foreclosure mortality table and the derived expectancy adjustment

TABLE V

FORECLOSURE EXPERIENCE WITH PROPERTY VALUATION ON SINGLE-FAMILY HOMES SECURING MORTGAGES INSURED, 1935–1939

Based on mortgages foreclosed through June 30, 1940

Property valuation	Mortgages insured 1935–1939 Number	Per cent	Foreclosure risk index, adjusted	Chi-square test of significance*	Foreclosure rate, adjusted 1935–1939
Less than $2,000	5,340	1.20	95.00	not significant	1.07%
$2,000 to 2,999	33,405	7.55	115.30	significant	1.31
3,000 to 3,999	81,036	18.41	82.35	significant	.93
4,000 to 4,999	97,889	22.08	81.70	significant	.93
5,000 to 5,999	85,424	19.27	77.01	significant	.87
6,000 to 6,999	61,617	13.87	101.95	not significant	1.15
7,000 to 7,999	30,457	6.87	110.39	not significant	1.25
8,000 to 9,999	26,082	5.88	176.53	significant	2.00
10,000 to 11,999	10,048	2.27	180.18	significant	2.04
12,000 to 14,999	6,373	1.44	238.80	significant	2.71
15,000 or more	5,144	1.16	213.70	significant	2.41
Total	443,377	100.00	100.00	significant	1.13%

TABLE VI

FORECLOSURE EXPERIENCE WITH BORROWER'S ANNUAL INCOME ON SINGLE-FAMILY OWNER-OCCUPIED HOMES SECURING MORTGAGES INSURED, 1935–1939

Based on mortgages foreclosed through June 30, 1940

Borrower's annual income	Mortgages insured 1935–1939 Number	Per cent	Foreclosure risk index, adjusted	Chi-square test of significance*	Foreclosure rate, adjusted 1935–1939
Less than $1,000	1,717	.47	120.70	not significant	1.30%
$1,000 to 1,499	19,632	5.40	102.00	not significant	1.11
1,500 to 1,999	68,724	19.01	97.09	not significant	1.05
2,000 to 2,499	90,300	25.01	80.44	significant	.90
2,500 to 2,999	54,075	14.99	82.15	significant	.88
3,000 to 3,499	41,840	11.58	99.31	not significant	1.07
3,500 to 3,999	27,488	7.61	116.32	significant	1.25
4,000 to 4,999	25,943	7.18	112.07	not very significant	1.21
5,000 to 6,999	19,001	5.26	140.30	significant	1.51
7,000 to 9,999	8,023	2.22	112.10	not significant	1.21
10,000 or more	4,716	1.30	170.15	significant	1.80
Total	361,425	100.00	100.00	significant	1.08%

* Significant-computed values of χ^2 are significant according to 1 per cent standard of probability.

Not very significant-computed values of χ^2 are significant according to 5 per cent standard of probability.

factors for mortgages exposed an average period of 12.13 months, 12.374 times as many, or 730 foreclosures, could have been expected had these 1939 mortgages been made in 1935.

Table IV, presenting the expectancy adjustment factors and the actual and adjusted foreclosures, illustrates how the total number of adjusted foreclosures is arrived at for all homes as well as for new and existing homes.

These expectancy adjustment factors derived from the mortgage mortality table were then applied as weighting factors to the numerical distributions of foreclosures on mortgages insured in each of the five years for each of the 16 risk characteristics selected for analysis. The adjusted foreclosures in each of the class intervals were summed to give the adjusted percentage distribution used in the calculation of the foreclosure relative or risk index.

It must be pointed out that the values of the foreclosure risk index so computed for any one of the risk elements are valid for only a five-year period of the order which describes the economic conditions which prevailed between 1935 and 1939.

In order to determine the reliability of the samples of cases used and the significance of the differences in the distributions of the mortgages-in-force and the foreclosures from assumed uniformity, the Chi-Square test of significance using the one per cent probability standard was applied.

Since the adjusted foreclosures are the number that could be expected if all the mortgages were exposed between 1 and 66 months, a *crude* foreclosure expectancy rate may be computed. At the end of five years of operation, the adjusted foreclosure rate on all mortgages *insured* by the Federal Housing Administration with comparable exposure periods is 1.12 per cent. This foreclosure rate is a crude rate and should be distinguished from the rates provided by the mortality table which are survivor rates.

Tables V and VI are taken from the FHA study of foreclosure experience to illustrate the application of this method of analysis in the case of property valuation and borrower's annual income.

SAMPLING ERRORS OF SYSTEMATIC AND RANDOM SURVEYS OF COVER-TYPE AREAS*

By James G. Osborne†
Forest Service

The purpose of the present paper is to report the results of an investigation into the accuracy of sample estimates obtained from systematic and from random samples. Also, in it will be described a procedure for estimating the sampling error of an estimate based on a set of systematic observations.

The studies reported here deal with the results of sampling to estimate the composition of an area by cover-type classes, but these estimates differ in no essential way from estimates of composition by many other criteria of classification, e.g., land-use classes, forest-condition classes, etc. In fact, the decision to use material of this kind in the study was based largely on convenience and availability of data, and not on a belief that the type of material selected is peculiarly suited to this kind of study. It should be pointed out, however, that the problem selected is one of real importance, since estimates of areas according to prescribed classifications are basic to most land-use planning and form an integral part of many surveys.

The data used in the studies reported here were obtained from two cover-type maps, one representing a moderate-sized area and the other a large area; the sampling being moderately intensive in the first and very extensive in the second. Specifically, the first is a vegetative-cover and land-use map of an area 28×30 miles in southern California; the sampling of this map being at the rate of one line per mile of width. The second is a forest-type and condition map of a section of northwestern Washington running southward 80 miles from the Canadian border, and from Puget Sound to the summit of the Cascades—a distance of roughly 60 miles; the sampling here was at the rate of one line in 10 miles of width, thereby following the national forest-survey procedure. In both cases, continuous lines were drawn across the maps and the length of line traversing each type was recorded.

The area in a specified type is assumed to occupy the same per cent of the total area as the length of line through that type does of the total

* A paper presented at the 103rd Annual Meeting of the American Statistical Association in joint session with the Institute of Mathematical Statistics, New York, December 30, 1941.

† The author is especially indebted to Miss M. M. Sandomire, who supervised and did much of the computational work as well as offered a number of fruitful suggestions in the analysis, and to Dr. I. T. Haig, for continued encouragement and support of the large clerical job involved.

length of line. As a check of this assumption, the area of cultivated land on the map of the smaller area was measured with a planimeter and was compared with that estimated by the line interception method, the estimate being based upon all lines measured. By planimeter measure, 41.748 per cent of the area was in cultivated land; by the line-interception method, the percentage was 41.896, a difference of 0.148 per cent

CHART I

A SYSTEMATIC SAMPLE OF 30 LINES ACROSS THE AREA IN CULTIVATED LAND, CONSISTING OF LINE 16 IN EACH OF THE 30 MILES OF WIDTH

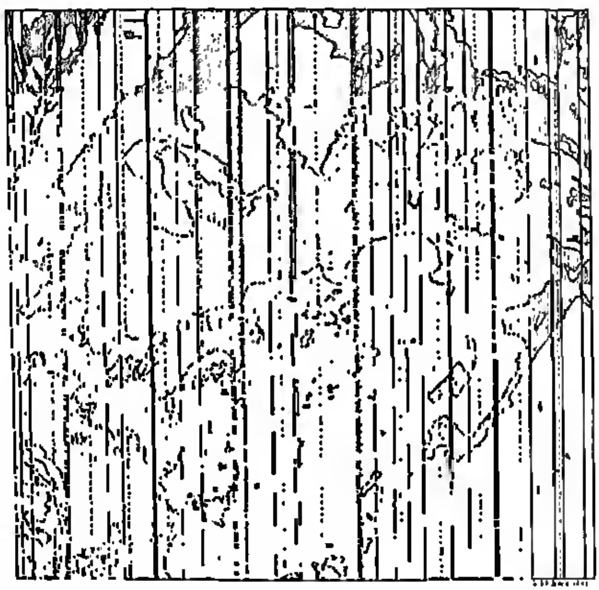
for the total area, or 0.353 per cent of the area in cultivated land. This is well within the limits of accuracy of the planimeter used.

The results of the studies of the two maps will be considered in turn. The data for the California study consisted of type recordings on 600 lines across the area, selected in the following way: Each of the 30 miles of the width of this tract was divided into 32 equal parts representing 32 lines across the area. From the numbers 1–32, 20 numbers were selected at random and the line corresponding to each of these 20 numbers was run in each mile. For example, line 4 was run in miles 1, 2, 3, . . . 30; then line 6 was run in each mile, etc. The observations on the 600 lines so selected constituted the basic data from which samples were drawn.

The 20 sets of 30 lines 1 mile apart constituted 20 systematic samples. They are, however, random in their totals and therefore provide a valid estimate of the variance of such complete surveys. For comparison, 20 random samples were drawn in each of several ways. These included among others (1) completely random (that is, sets of 30 lines selected completely at random out of the 600 lines) and (2) stratified

CHART II

A RANDOMIZED BLOCK SAMPLE OF 30 LINES ACROSS THE AREA, CONSISTING OF 2 LINES SELECTED AT RANDOM IN EACH OF 15 BLOCKS 2 MILES WIDE



random or randomized blocks. In (2), the area was divided into 15 blocks, each 2 miles wide, and two lines were drawn at random in each block. Time permits reporting upon only one cover type, and for this purpose the area in cultivated land is selected. For the other types the implications are similar.

Charts I and II show the area in cultivated land (cross-hatched), and two methods of sampling.

To the nearest whole number, the standard deviation of the systematic totals was 279 chains; that of the randomized block totals was 564; and that of the completely random totals was 1701. The means of the survey totals were, respectively, 28164, 28238, and 28657. It is thus seen that, for this type, the estimates by all three methods tend toward the

same value, but that the standard deviation of the systematic totals is only half as large as that of the randomized blocks, and only one-sixth as large as for completely random samples.

The practical significance of the comparison of the accuracy of systematic and randomized block estimates is impressive. Only one-fourth as much information is obtained from the stratified random as from the systematic sample. Why? At least a part of the answer is apparent from a glance at the distribution of the observations composing the first randomized block sample drawn (see Chart II). As would be expected with blocks 2 miles wide, approximately one-fourth of the miles have no samples (actually 7 out of 30) while the same number have 2 lines. Even more serious, as will generally be true, lines close together occur frequently, there being in this case 4 pairs of lines only $\frac{1}{8}$ mile apart. With lines 28 miles long, a second line $\frac{1}{8}$ mile from the first gives little additional information after the first has been measured. To offset this weakness, an unbiased estimate of the sampling error of a survey of this type is directly and easily calculated by a straightforward application of random sample error formulae. An estimate of the standard deviation among whole surveys, of 575, calculated from the average within-sample mean-square, is close to the estimate of 564 based on survey totals.

Considering the systematic surveys again, it is well to examine the nature of the variation found in populations in place. It will be recognized first that subdivision of an area into blocks does not divide it into homogeneous strata. Rather, a variate changes continuously within a block, and from block to block. Generally, then, a variate measured first at a particular place and then at a place a differential distance away will undergo a differential change. The variate may, then, be considered as a continuous function of position and the problem of sampling reduces, sensibly, to one of curve fitting. Further, the correlation between observed values depends upon the distance between the points of observation.

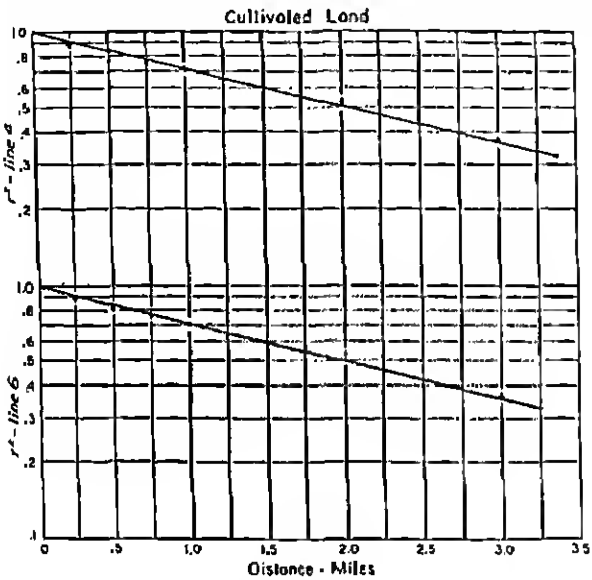
As with randomized block sampling, it is evident that since each mile is represented in the sample to the same extent as in the total area, any differences between miles contribute nothing to the sampling error. Also, it may be said that the observation in any mile is an estimate of any other observation in that mile. The problem may be restated, then, as: "Given the results of a survey, what is the error of predicting the results of another similar survey?" Considered in the light of linear least-squares regression theory, the measured length of line through a type, in any mile, may be considered the independent variable, and the

estimate for any other line (not measured), in the same mile, as the dependent variable. Then the variance of estimating any other line is the variance of estimating its value when only the mile within which it lies is known, multiplied by $(1-r^2)$, where r^2 is the square of the correlation coefficient of the line measured and the one to be estimated. The

CHART III

CORRELATION COEFFICIENTS OF LINES 1, 2, AND 3 MILES APART BASED ON LINES NOS. 4 AND 6

Correlation of observations on lines 4 and 6 with that on other lines less than 1 mile distance also are shown.



correlation coefficient depends, for its value, upon the distance between the measured and estimated line.

This correlation coefficient can be estimated by calculating the correlation of lines measured at distances of one unit (mile in this case), two units, etc., and plotting the correlation coefficient as the ordinate and the distance as the abscissa. As a control in drawing the curve, it is known that at zero distance the correlation coefficient is one. Experience indicates the relationships to be exponential in form, hence plotting on semi-logarithmic paper is helpful.

Chart III shows the calculated correlation coefficient of lines 1, 2, and 3 miles apart for systematic samples, lines Nos. 4 and 6 in every mile being the bases of the two parts of the figure. Values of the corre-

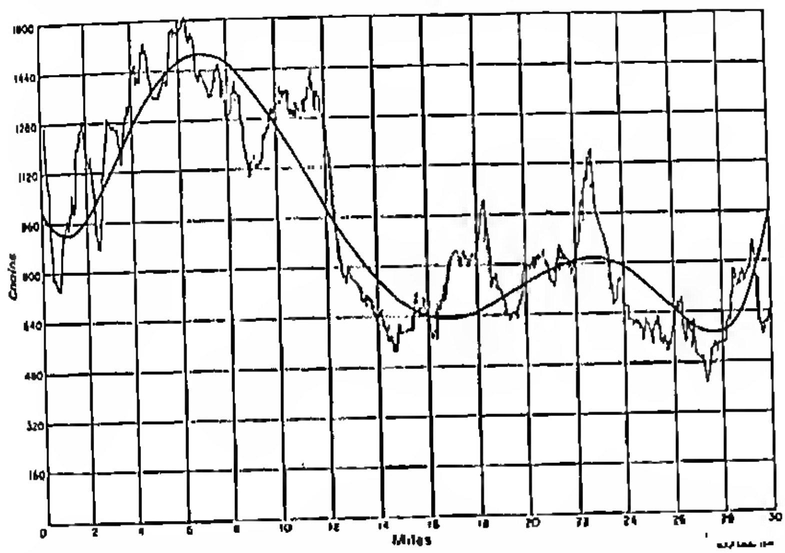
lation coefficient for lines less than one mile apart also are calculable in this study and are shown here.

The variance of an observation at a single place is taken as the residual mean-squared deviation from a polynomial fitted by least squares. Chart IV shows a plot of the observations of the 600 lines measured, the number of chains (66 feet) of cultivated land crossed by the line

CHART IV

NUMBER OF CHAINS (66 FEET) OF CULTIVATED LAND CROSSED BY EACH OF 600 LINES PLOTTED OVER THE POSITION OF THE LINE (IN MILES FROM THE LEFT EDGE OF THE MAP)

A polynomial of degree 6 has been fitted to the observations on line No. 16 in every mile by the method of least squares. Observations used are shown with heavier dots.



being plotted on the position of the line in miles from the left edge of the map. A polynomial of degree 6 has been fitted to the observations on line No. 16 in every mile.

Experience has shown also that, when enough terms in the polynomial have been used so that the introduction of terms of higher degree does not reduce the residual mean square significantly at the 5 per cent level, the residual mean square is approximately equal to half the mean-squared successive difference of the original observations; i.e., $\delta^2/2$, where

$$\delta^2 = \frac{\sum_{i=1}^{n-1} (X_{i+1} - X_i)^2}{n - 1}.$$

In this equation, the X_i are the observations in miles 1, 2, 3, . . . , $n-1$, n, and n is the number of lines in the sample.

The average squared error of estimating any survey from any observed survey is, then,

$$S_T^2 = ns^2(1 - \bar{r}^2),$$

where s^2 = residual mean square from a least-squares polynomial,
n = the number of lines in a survey,
and $\bar{r}^2$ = the estimated average of the squares of the correlation coefficients of a measured line with all lines within the mile within which the observation occurs. For example, if line 10 is observed, $\bar{r}^2$ is the average squared correlation coefficient for lines between 0 and 10/32 mile and between 0 and 22/32 mile. In practice this is most easily obtained by integrating under the curve $r^2 = e^{-kd}$ between the limits 0 and 10/32, and, 0 and 22/32. In this equation d is the distance between lines and k is a constant, equal to 2 log $_e r_{(1)}$ where $r_{(1)}$ is the correlation coefficient of lines one unit apart.

TABLE I

SUMMARY OF ESTIMATES OF STANDARD DEVIATIONS OF SAMPLE TOTALS FOR FOUR MAJOR TYPES CALCULATED BY METHODS DESCRIBED ABOVE

Type	Systematic				Randomized block		Completely random	
	$S.D._T$*	$\sqrt{ns^2(1-r^2)}$†	$\sqrt{n\frac{\delta^2}{2}(1-r^2)}$‡	$S.D._T$§	$S.D._T$‖	$S.D._W$¶	$S.D._T$**	$S.D._W$††
	Chains	Chains	Chains	Chains	Chains	Chains	Chains	Chains
Cultivated	270	254	240	004	561	575	1701	1731
Shrub	202	311	272	005	470	514	1610	1685
Grass	103	154	152	310	203	257	418	400
Woodland	141	130	133	217	166	188	224	254

* Standard deviation of totals of 20 randomly selected systematic samples.

† Square root of average of estimates of variances of totals based on systematic sample observations, estimated from the residual mean square from a polynomial fitted by least squares and the estimated average squared correlation coefficient of observed survey with all possible systematic surveys.

‡ Same as † with half the mean-squared successive difference ($\delta^2/2$) substituted for the residual mean square from a fitted polynomial.

§ Square root of average variance of totals estimated on *assumption* that lines systematically 1 mile apart were randomly located, two in each of fifteen 2-mile-wide blocks.

‖ Standard deviation of 20 sample totals, the samples consisting of 2 lines at random in each of 15 blocks.

¶ Estimate of standard deviation of sample totals (sampling according to ‖) based upon the average within-sample mean square.

** Standard deviation of 20 sample totals, the samples consisting each of 30 completely randomly selected lines.

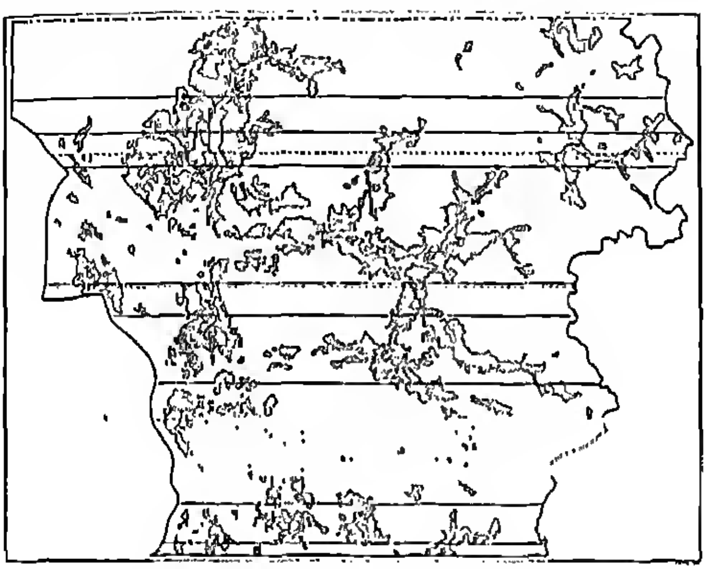
†† Estimate of ** based upon the average within-sample mean square.

Upon applying this formula to each of the systematic surveys, the standard deviations were found to range from 192 to 308, the standard deviation corresponding to the average variance being 254. This is to be compared with 279, the standard deviation among the survey totals. By comparison, the standard deviations of totals from the individual

CHART V

MAP SHOWING AREA IN DOUGLAS-FIR

One of the 30 randomized block samples consisting of 2 lines selected at random in each of four 20-mile-wide blocks is shown. The block boundaries are indicated by the broken lines.



randomized block samples ranged from 431 to 770, that corresponding to the average variance being 575, as mentioned earlier.

Values of "t" were calculated for each of the systematic surveys and a comparison was made with the theoretical distribution of "t." A Chi-square value of 10.5 was found with 10 degrees of freedom.

Table I summarizes the results for the four major types found in the area.

For the Washington map, only the results for the Douglas-fir type will be mentioned. Here, totals from 30 sets of 8 lines uniformly 10 miles apart showed a standard deviation of 785 chains. Randomized block samples based on 2 lines in each of four 20-mile-wide blocks showed a standard deviation of 1,002 chains. Chart V shows the distribution of the Douglas-fir type in the area studied with one of 30

randomized block samples observed. The variance of the random surveys is seen to be about twice (i.e., 1.83 times) as large as for systematic surveys of the same *intensity*. Again, the mean values were very close to being 8,340 and 8,280 chains for the systematic and the randomized block surveys, respectively.

With the aid of the more simply calculated formula,

$$S.D._{T} = 8\frac{\delta^2}{2}(1-r^2)$$

(where δ^2 and r^2 have the same meaning as previously), the estimated standard deviations of totals were calculated from each of the 20 systematic surveys. These ranged from 342 to 1,544 with that derived from the average variance being 955. This is to be compared with 785 chains, the standard deviation of the actual totals.

To summarize these tests of stratified random and systematic surveys of cover-type areas:

1. For the material used in the test, stratified random surveys were only one-half to one-fourth as efficient as systematic surveys of the same intensity;

2. If data taken systematically are used with random sample formulae, biased estimates of the sampling errors of totals or means result;

3. It is demonstrated that with material of this kind, random sample formulae when applied to randomly selected observations yield dependable estimates of the sampling errors of totals; and

4. It is demonstrated that from estimates of the correlation of measured and unmeasured lines dependable estimates of the sampling errors of systematic samples are obtained.

SAMPLING WITH TRANSVERSE TRAVERSE LINES

By Malcolm J. Proudfoot*
Bureau of the Census

In the analysis of land use for the Tennessee Valley Authority, a rapid reconnaissance technique was needed to sample the quantity and distribution of various types of land at lower cost than the expense of detailed mapping. In the years 1934–35, a traverse method of estimation was tried out with the purpose of determining the adequacy of sampling data so collected.

Traverses spaced at various intervals were run across completed land-use maps and the lengths of various types of land were noted along each traverse. These lengths, converted to percentages, were compared with the percentages based on the areas taken by these types of land as determined by planimeter measurement. This is essentially the method introduced by Rosiwal[1] in 1898 for petrographic analysis. He computed the content of various mineral components of rocks by measuring the lengths occupied by these components along traverse lines. J. M. Trefethen of the University of Wisconsin first applied Rosiwal's method to geographical reconnaissance.[2] His conclusions coincided with Rosiwal's, namely, that good results could be obtained when the total length of the traverse lines exceeds 100 times the average intercept of the field types traversed.

The importance of this ratio is illustrated by Chart I. Here is shown a large square divided into fields of different types labelled from A to E. Table I A, pertaining to the entire area assumed to be sixteen square miles, gives in Column (1) the percentages taken by each field type as determined by planimeter measurement, and in Column (2) the percentages taken by each field type as determined by measuring the total traverse intercepts shown by broken lines. Column (3) shows the deviations between Columns (1) and (2) which result in a root-mean-square error of 0.6. In this case the ratio of the total traverse length to average field intercept is 88 to 1.

In contrast, Table I B, pertaining to the area of four square miles (the lower right-hand quarter of Chart I), sampled by the same trav-

* The author acknowledges with gratitude several helpful suggestions made by Dr. W. Edwards Deming, Staff Mathematician of the Bureau of the Census, and the invaluable assistance rendered by Louise Wahlruff in these arduous statistical tasks.

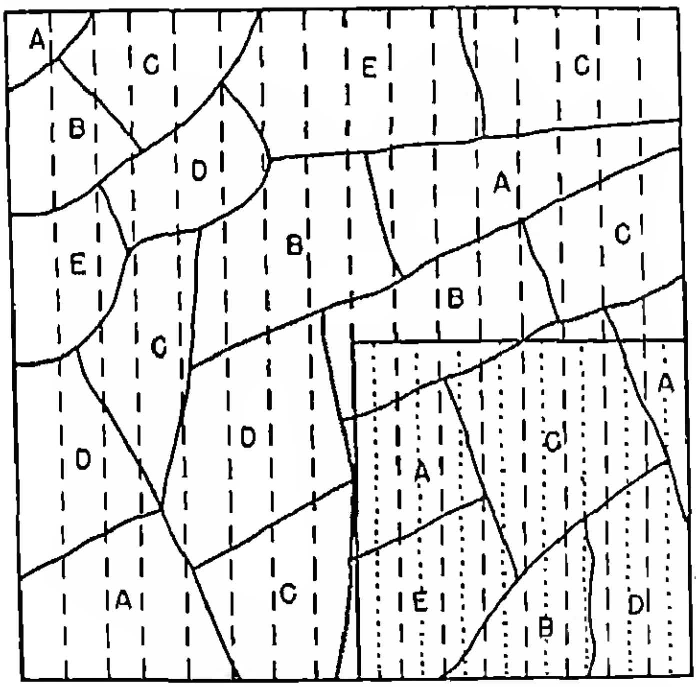
[1] August Rosiwal, *Ueber geometrische Gesteinsanalysen, Ein einfacher Weg zur Ziffermaessigen Feststellung des Quantitaetsverhaeltnisses der Mineralbestandtheile gemengter Gesteine* (Wien: Koenigliches Kaiserliches Reichsanstalt, 1898). For a short summary of this work in English, see Albert Johannsen, *Manual of Petrographic Methods* (McGraw-Hill Book Co., 1918), pp. 291–299.

[2] Paper read before the Association of American Geographers, St. Louis, Missouri, December 1935.

erse spacing shows nearly a fourfold increase in the root-mean-square error. This is accompanied by the less refined ratio of 23 to 1 although it will be noted by inspection that the average field intercept is approximately the same as for the entire area.

A higher ratio achieved for the same area by twice the frequency of lines (the additional lines are shown by dots) gives the results shown in Table I C. The ratio of total traverse length to average field intercept

CHART 1



is now 46 to 1 and the root-mean-square error has decreased approximately 50 per cent.

From these comparisons it seems obvious that the greater the total traverse length and the narrower the spacing the smaller the resulting root-mean-square error in sampling the five field types. The root-mean-square errors obtained for these three examples, when plotted, give a suggestion of regularity. At infinitely close spacing there would be no error at all, hence the line must pass through the origin.

The percentages of the total areas taken by the five field types contributing to the root-mean-square error do not differ widely; none pre-

dominates over the other and none is extremely small. It remains to see what happens to the error of the traverse method if the percentage taken by one of the field types is small.

To obtain evidence on this question and in an effort to work out a traverse precision chart for reconnaissance field purposes, the ade-

TABLE I

COMPARISON OF TRAVERSE LENGTH WITH PLANIMETER MEASUREMENT

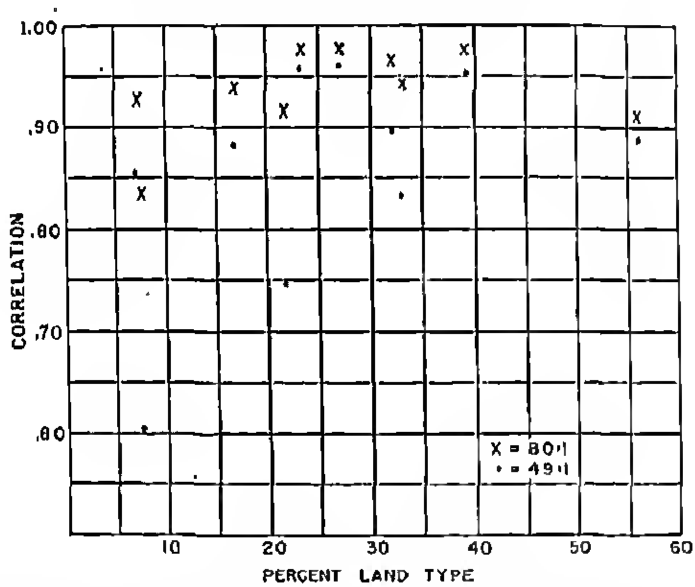
Field type	Planimetered field measurements in per cent (1)	Traverse line measurements in per cent (2)	Deviation in per cent (3)
A. Ratio of Total Traverse Length to Average Field Intercept is 88:1			
A	18.3	19.2	0.9
B	17.3	16.8	0.5
C	30.8	31.3	0.5
D	18.3	18.0	0.3
E	15.3	14.7	0.6
		Root-mean-square error	0.6
B. Ratio of Total Traverse Length to Average Field Intercept is 89:1			
A	20.5	22.0	1.5
B	17.3	15.3	2.0
C	28.5	28.4	0.1
D	16.0	19.3	3.3
E	17.7	15.0	2.7
		Root-mean-square error	2.2
C. Ratio of Total Traverse Length to Average Field Intercept is 40:1			
A	20.5	19.6	0.9
B	17.3	18.4	1.1
C	28.5	28.6	0.1
D	16.0	14.4	1.6
E	17.7	19.0	1.3
		Root-mean-square error	1.2

quacy of half- and quarter-mile spacings were tested out on an area comprising 134 square miles and ten different land types. The area stretched from the Holston River to the Smoky Mountains and the various land types ranged from 56 per cent down to 7 per cent of the total area. In carrying out the experiment, the total area was divided into 134 square miles, and the percentage of each square devoted to each of the ten land types was computed from their areas as obtained by planimeter measurement. The results obtained by the half- and quarter-mile traverse lines, likewise expressed by percentages devoted to each land type, were then compared with the percentages obtained by the planimeter. For any land type in the entire field there were thus 134 percentages based on planimeter measurements, and 134 percent-

ages based on the traverse lines. These 134 pairs were plotted on a scatter diagram, for which a coefficient of correlation was computed. The results are shown in Table II and Chart II. The correlations are adequate, it will be observed, except for the land types having the lowest percentage. The crosses (showing the correlations for the 80 to 1 ratios) invariably fall above the dots (showing the correlations for the 40 to 1 ratios), and moreover are more nearly uniform for the ten different percentages.

CHART II

CORRELATIONS BETWEEN TRAVERSE AND FULL COVERAGE MAPPING FOR TEN DIFFERENT LAND TYPES



These coefficients of correlation show that if the ratio of the total length of the traverse line to the average field intercept of all fields is sufficiently large, adequate results can be obtained for a field type comprising any given percentage of the total area. However, for a field type comprising less than 5 per cent, this ratio apparently must be increased beyond 100 to 1, at which point the traverse field costs approach the cost of full coverage. These indices therefore provide at least the beginning of a traverse precision chart for use in forecasting the accuracy of traverse sampling under varying field conditions.[3]

[3] Had funds been available this chart might have been greatly extended and refined. As observed by Dr. Deming, it would have been highly desirable to study the reliability of the results under different ratios, using the average field intercept of each field type to the total traverse length. It is hoped that someone with similar data and the funds and facilities will undertake such investigations.

To show the utility of this chart, let us assume an area of approximately 100 square miles. Limited funds and time render detail mapping impracticable, yet quantitative data pertaining to the proportional

TABLE II

CORRELATIONS BETWEEN TRAVERSE AND FULL COVERAGE MAPPING FOR TEN DIFFERENT LAND TYPES

Type of land	Arithmetic mean of the detailed field data in per cent (1)	Arithmetic mean of the traverse data in per cent (2)	Coefficients of correlation between the traverse data and the detailed field data (3)
	(Half-mile traverse spacing)		
Not eroding...........	56.1	58.1	.8873
Woodland.............	30.3	30.8	.0507
Sheet erosion..........	33.3	32.8	.8323
Cropped.............	32.2	31.1	.8052
10–20 per cent slopes...	27.1	20.0	.0578
0–10 per cent slopes....	23.3	23.8	.0501
Pastured..............	21.6	21.8	.7452
20–40 per cent slopes...	10.8	17.0	.8832
Idle..................	7.7	8.1	.0024
Gullying..............	7.4	7.0	.8542
	Total traverse length / Average field intercept	40	
	(Quarter-mile traverse spacing)		
Not eroding...........	56.1	57.0	.0127
Woodland.............	30.3	30.2	.0707
Sheet erosion..........	33.3	33.1	.0401
Cropped.............	32.2	31.0	.0057
10–20 per cent slopes...	27.1	27.2	.0706
0–10 per cent slopes....	23.3	23.0	.0701
Pastured..............	21.6	21.0	.0107
20–40 per cent slopes...	10.8	17.7	.0376
Idle..................	7.7	7.0	.8322
Gullying..............	7.4	7.0	.0253
	Total traverse length / Average field intercept	80	

distribution of the major land types in this area are required. The distribution of these types is to be shown by square-mile units. To meet this objective, the following 10 steps can be taken:

1. Map three widely, but evenly, spaced traverse lines across the area. If possible, run these lines transverse to the drainage pattern. From the field data thus obtained compute the average field intercept

of all fields and the percentage of the total area taken by each of the major land types.

2. Compute the ratio of the total traverse length to the average field intercept. Refer to the traverse precision chart. If this ratio is sufficiently high, a high coefficient of correlation is indicated for each land type comprising more than 7 per cent of the total sample. If these conditions are met the traverse lines adequately sample the entire area for the major land types concerning which information is desired.

3. Lay off a square-mile grid on a master base map of the area. If possible, place this grid so that one side or the other of each square mile lies transverse to the drainage pattern of the area.

4. Using the determined average field intercept and the percentage figures for the major land types, determine from the traverse precision chart the traverse spacing which will provide the necessary ratio within each square mile to give satisfactory coefficients of correlation.

5. Lay out the determined traverse spacing on the grid of the master base map.

6. Transfer these traverse lines to field maps cut to an appropriate note-book size.

7. Map the land types occurring along each traverse line.

8. By means of the traverse field data thus obtained compute the percentage of each land type occurring within each square mile.

9. Plot these percentage figures on the square-mile grid of an outline map for each land type.

10. Draw isopleth lines of equal percentage for each land type, selecting intervals suited to the purpose of the investigation.

INDEX-NUMBER DIFFERENCES: GEOMETRIC MEANS

By Irving H. Siegel
U. S. Bureau of Labor Statistics

This is the last of three papers in this Journal[1] on the difference between indexes obtainable from the same set of relatives. The first two papers were concerned with the difference between the Paasche and Laspeyres formulas and the more general case of the difference between any two arithmetic means. The area covered in those papers is rather broad, since any harmonic mean of relatives and any weighted or unweighted aggregative measure, including the Edgeworth, can also be expressed as an arithmetic mean of relatives.[2] With the extension of the discussion here to geometric means of relatives and to the "ideal" index, the exploration of the entire domain of practical index numbers is virtually completed.

The methods employed in this paper will be similar to those employed in the earlier ones. In addition, use will be made of the well-known facts that logarithms of geometric means reduce to linear forms and that arithmetic means are necessarily greater than similarly weighted geometric means if not all of the relatives are equal.[3] The use of logarithms will make it most convenient for us to compare the magnitudes of the different index numbers in ratio form.

First, we shall consider the ratio (R) between weighted and unweighted geometric means (G_w and G, respectively) of the same set of n relatives ($X_i = q_i'/q_i$):

$$R = G_w/G = (\Pi X^w)^{1/\Sigma w}/(\Pi X)^{1/n}.$$

Since

$$\log R = \frac{\sum w \log X}{\sum w} - \frac{\sum \log X}{n} = \frac{\begin{vmatrix} \sum w \log X & \sum w \\ \sum \log X & n \end{vmatrix}}{n \sum w}$$

and

$$\left\| \begin{matrix} \sum w \log X & \sum w \\ \sum \log X & n \end{matrix} \right\| = \left\| \begin{matrix} w_1 & \cdots & w_n \\ 1 & \cdots & 1 \end{matrix} \right\| \cdot \left\| \begin{matrix} \log X_1 & \cdots & \log X_n \\ 1 & \cdots & 1 \end{matrix} \right\|,$$

[1] See Irving H. Siegel, "The Difference between the Paasche and Laspeyres Index-Number Formulas," this Journal, September 1941, pp. 343–350, and "Further Notes on the Difference between Index-Number Formulas," December 1941, pp. 519–521.

[2] It should be noted that unweighted arithmetic and harmonic means can also be written as weighted arithmetic means: $\Sigma X/n = \Sigma wX/\Sigma w$, where $w_i = 1 = w_i'(1/w_i')$, and $n/\Sigma(1/X) = \Sigma kX/\Sigma k$, where $k_i = 1/X_i$.

[3] Various algebraic devices leading to approximate expressions (e.g., the expansion of logarithms into series) could also have been used.

we have

$$\log R = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \begin{vmatrix} w_i & w_j \\ 1 & 1 \end{vmatrix} \cdot \begin{vmatrix} \log X_i & \log X_j \\ 1 & 1 \end{vmatrix}}{n \sum w}.$$

From the numerator of the last expression (which contains ${}_nC_2$ determinant products), it is obvious that $G_w > G$ necessarily if the rank coefficient of correlation between the w_i and the $\log X_i$ equals $+1$, and that $G_w < G$ necessarily if the coefficient equals -1. It can also be shown that, since

$$\log R = r_{w \log X}\, \sigma_w \sigma_{\log X}\, n / \sum w = f(r),$$

or

$$G_w = 10^{f(r)} G,$$

we have $G_w \gtreqless G$ according as $r_{w \cdot \log X} \gtreqless 0$. These criteria[4] are analogous to those derived for the difference between a weighted arithmetic mean and an unweighted one.

We shall now consider the relation between two differently weighted geometric means:[5]

$$R' = G_{w'}/G_w = (\Pi X^{w'})^{1/\Sigma w'} / (\Pi X^{w})^{1/\Sigma w}.$$

Since

$$\log R' = \frac{\sum w' \log X}{\sum w'} - \frac{\sum w \log X}{\sum w} = \frac{\begin{vmatrix} \sum w' \log X & \sum w' \\ \sum w \log X & \sum w \end{vmatrix}}{\sum w \sum w'}$$

and

$$\left\| \begin{matrix} \sum w' \log x & \sum w' \\ \sum w \log x & \sum w \end{matrix} \right\| = \left\| \begin{matrix} w_1' & \cdots & w_n' \\ w_1 & \cdots & w_n \end{matrix} \right\| \cdot \left\| \begin{matrix} \log X_1 & \cdots & \log X_n \\ 1 & \cdots & 1 \end{matrix} \right\|,$$

we have

$$\log R' = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \begin{vmatrix} w_i' & w_j' \\ w_i & w_j \end{vmatrix} \cdot \begin{vmatrix} \log X_i & \log X_j \\ 1 & 1 \end{vmatrix}}{\sum w \sum w'},$$

[4] Since an unweighted arithmetic mean may be written as a weighted one (see footnote 2), $\log R$ may also be expressed in terms of a weighted correlation coefficient and weighted standard deviations.

[5] W. V. Lovitt has investigated the sign of the difference between many index-number formulas, including geometric means (see *Cowles Commission: Report of Third Annual Research Conference*, 1937, pp. 85–87). His investigation was restricted, however, to indexes with the four sets of weights commonly applied to price relatives. Our discussion is more general and embraces Lovitt's results as special cases.

Incidentally, an arithmetic and a geometric mean with the *same* set of weights are improperly included in Lovitt's list of index pairs for which the inequality sign is "not permanently directed." (An arithmetic and harmonic mean with the same set of weights are also improperly included.)

which may also be written as

$$\log R' = r_{w:(w'/w)\cdot \log X}\,\sigma_{w:(w'/w)}\,\sigma_{w:\log X} \sum w / \sum w' = f(r').$$

From the last two expressions, it is evident that $G_{w'} > G_w$ necessarily if the rank coefficient of correlation between the w_i'/w_i and the $\log X_i$ equals $+1$; that $G_w' < G_w$ necessarily if the coefficient equals -1; and, more generally, that the sign of the difference between the two indexes corresponds to the sign of the weighted correlation coefficient $r_{w:(w'/w)\cdot \log X}$. These criteria are analogous to those shown in the earlier papers for the difference between two arithmetic means with dissimilar weights.

Although an arithmetic mean is greater than a similarly weighted geometric mean (if some of the relatives are not equal), it is not necessarily greater than a geometric mean with different weights.[6] Designating the arithmetic mean by M_w, the geometric mean by G_w', and the ratio by R'', we have

$$R'' = M_w/G_{w'} = \frac{\sum wX}{\sum w} \Big/ (\Pi X^{w'})^{1/\Sigma w'},$$

whence

$$\log R'' = \log\left(\frac{\sum wX}{\sum w}\right) - \frac{\sum w \log X}{\sum w}$$
$$- r_{w:(w'/w)\cdot \log X}\,\sigma_{w:(w'/w)}\,\sigma_{w:\log X} \sum w / \sum w',$$

or

$$R'' = M_w/G_w \cdot 10^{f(r')}.$$

Since $\log(\Sigma wX/\Sigma w) > \Sigma w \log X/\Sigma w$ (if the X_i are not all equal), it is clear that $R'' > 1$ (i.e., $M_w > G_w'$) if $r_{w:(w'/w)\cdot \log X} \leqq 0$. A positive correlation coefficient, however, is consistent with values of R'' greater or less than unity.

Like its components, the Paasche and Laspeyres indexes, the "ideal" index need not be greater than a weighted geometric mean of relatives. We first consider the relation between the "ideal" index, $I = (P_X L_X)^{1/2}$, and $G_m = (\Pi X^m)^{1/\Sigma m}$; P_X and L_X represent the Paasche and Laspeyres indexes of the $X_i = g_i'/g_i$, and the m_i are the weights in both the

[6] Thus, an "unadjusted" industry production index of the kind shown by S. Fabricant, *Output of Manufacturing Industries: 1899–1937* (1940), need not exceed an index computed for the same products by the Day-Thomas formula. The Fabricant measures are based on the Edgeworth formula; the Day-Thomas index for the time t_i is a geometric mean of the production relatives weighted by averages of the money values for t_i and the base period, t_0.

Laspeyres index and the geometric mean. If the X_i are not all equal, it may readily be shown that $I > G_m$ necessarily when $P_X \geqq L_X$, or, more generally, when $r_{m:Y \cdot \log X} \geqq 0$. When the correlation coefficient is negative, the inequality sign between the two indexes is not definitely directed. If, instead of G_m, we take $G_{mY} = (\Pi X^{mY})^{1/\Sigma mY}$ (where the m_iY_i are the weights in P_X), then $I > G_{mY}$ when $L_X \geqq P_X$, or, more generally, when $r_{m:Y \cdot \log X} \leqq 0$. When the correlation coefficient is positive, the inequality sign between the indexes is not definitely directed. If, finally, we take $G_w = (\Pi X^w)^{1/\Sigma w}$, then $I > G_w$ necessarily when both $r_{mY:(w/mY) \cdot \log X}$ and $r_{m:(w/m) \cdot \log X}$ are zero or negative.

The difference between the "ideal" index and other measures considered in the earlier papers requires little comment. Its relation to its components, the Paasche and Laspeyres indexes, is obvious; and Professor Fisher has conclusively shown how close it is to the Edgeworth index.[7] Other cases may readily be investigated by the methods employed here and in the earlier papers.

[7] See I. Fisher, *Making of Index Numbers* (1922), pp. 428–430; H. T. Davis and W. F. C. Nelson, *Elements of Statistics* (1937), pp. 111–112; and H. T. Davis, *Theory of Econometrics* (1941), p. 328.

MATHEMATICAL OPERATIONS WITH PUNCHED CARDS*

By J. C. McPherson
International Business Machines Corporation

In the past fifty years we have seen a very significant change in the extent and complexity of computations required to apply mathematical formulae to concrete situations. The early laws of physics, mechanics and chemistry were expressed very frequently by mathematical expressions for which the computations were fairly simple, and well within the range of the mechanical devices, including tables, then available.

More recently we have been faced, as higher branches of mathematics have been called into use in solving other phenomena, with an immense increase in the amount of labor required to compute results under the mathematical expressions developed for this work. As an example, we might mention boundary value problems and the use of determinants in analyzing statistical correlations and for solving electrical networks. While the mathematical expressions are simple, the actual labor of carrying out the computations indicated can be a matter of weeks of work.

There thus arises a further course for mathematical study, the development of mathematical expressions or expansions whose computation can be effected by the means at our disposal. This involves determining the relative simplicity *in use* of various mathematical forms and the establishment of additional information regarding the accuracy and limits of error of various processes which can be carried out by the devices which we now have and can project.

It is my purpose in this paper to describe briefly one of the more powerful but little used tools for extensive mathematical computation in order to point out its general function and its present application to computing problems.

Punched-card tabulating machines, because they require no further manual work than the original entry of the problem on punched cards which afterwards actuate automatic machines, form the most powerful tool yet devised for the performance of mathematical computations. As yet the full capabilities of the automatic punched-card method have not been achieved in scientific fields, except in isolated instances in widely separated fields of activity. In bringing before you the results of the use of machines in the various fields of science, it is hoped that many addi-

* A paper presented at the 103rd Annual Meeting of the American Statistical Association in joint session with the Institute of Mathematical Statistics, New York, December 28, 1941.

tional uses for machines will be developed and that thinking in machine terms will be greatly stimulated.

Electrical punched-card accounting machines were developed primarily as a result of the analytical demands of the census and their great growth has been due to their usefulness in handling statistical and accounting procedures. This has led to the development of a specific series of machines actuated by the punched-cards on a functional basis, i.e., each machine designed to perform a specific function and record its results in form for further automatic machine handling if desired.

The basic feature of tabulating equipment is its ability to read punched holes and perform the computations indicated by the holes, recording the results in printed or punched-hole form for subsequent processing. These machines read a line at a time and with automatic reading of the cards goes a high speed of computing and handling of individual problems. Each machine has the ability to handle tabulating cards regardless of arrangement of data.

There are some six punched-card machines whose usefulness in handling mathematical work has been demonstrated. For computing work, the automatic Multiplier is perhaps the most useful. This machine reads multiplication problems, performs the computations, and punches the answer back in the card on which the problem is stated. The Multiplier can also make cross additions or subtractions while multiplying, thus performing in a single step such operations as linear interpolation. Its operation is completely automatic and at a speed several times that of a clerk with a computing machine. On such work as 8 by 8 multiplications, machine speed is 750 multiplications per hour and on smaller problems speeds up to 1,500 are obtained.

Of next importance is the machine termed the Reproducer. This machine can transfer all or part of the punched information on one card or set of cards to another set of cards at the fixed speed of 100 cards per minute. It is used, for example, in making copies of punched-card tables or parts of tables; for combining intermediate results computed on different sets of cards onto a single card for further processing; and for transferring data from one set of cards to another. The Reproducer is unique in its ability to copy information from one document to another. At 6,000 cards per hour we are able to reproduce, rearrange, or extract information from punched-card records.

An automatic Sorting Machine is available for rapidly rearranging a set of cards into another sequence or for bringing together all cards carrying a similar punched-card designation. It operates at 400 cards per minute.

Another card arranging machine is the Collator, which can interleave two separate files of cards into a single file; or select cards from the one file matching cards in the other file; or select all cards greater than a certain value, or less than a certain value, or falling between specified limits. This machine is used to select cards from a table file, and to refile the selected cards after use.

The principal punched-card machine is called the Electric Accounting Machine. It is a giant printing adding machine actuated by the passage of the cards through a card feed at speeds as high as 150 cards a minute. It has the ability to add, subtract, or eliminate amounts punched in one or several fields of the cards passing through it. It automatically adds all cards having a common designation; and at the end of the group, which it determines automatically, it prints the total and punches a new card with the group designation and the group total on an automatic punch electrically connected to the accounting machine. The machine has a maximum adding capacity of 80 digits. These adding wheels may be grouped at will into counters of varying size and several factors may be added simultaneously. This machine is the commercial version of Babbage's "Differential Engine" capable of operation over any number of orders of differences, and a counter large enough to handle figures of practically any required size.

There are several kinds of punches available for originally recording data on the punched cards, and they are designed for rapid operation. There are also Verifiers for checking punching by a second recording of the data and Interpreters for printing on the cards the information punched in the card.

The use of the punched-card method for mathematical computation involves the use of one or more of the machines briefly described above. Some of the techniques require only the use of the Punches, Sorter and the Electric Accounting Machine, while other operations will call for the use at some point of all of the machines described above.

Tables. An extremely important use for the punched-card equipment is the preparation of tables. The equipment is so powerful in this respect that it has been said that every computational problem should be examined to see to what extent special tables prepared by machine can be used in its solution. This statement applies not only to processes conducted entirely by the machine method, but also to extensive problems where the special tables can be prepared by the machine and then used by computers in their further work. Even in so simple a thing as making a linear table, in one instance the first thousand multiples of each of four 20-figure numbers were produced in 4 hours, i.e., in about

one-tenth of the time in which they could have been copied by hand.

The Electric Accounting Machine and Summary Punch are used in the preparation of tables using the method of differences. The process is covered in detail in a recent paper "On the Mechanical Tabulation of Polynomials," appearing in the September, 1941, issue of the *Annals of Mathematical Statistics.*

Interpolation of tables is frequently performed by aid of punched-card techniques. On systematic interpolation, a paper by L. J. Comrie in 1928 on "Construction of Tables by Interpolation" explains the preparation of subdivided tables by computing the last digit of each interpolated value exactly by the aid of a set of prepunched cards. These figures are then differenced until they are smooth and the entire value of the differences inferred. From these differences, the Electric Accounting Machine will then automatically construct the subdivided values of the function.

Interpolation by use of the Lagrangian formulas can be readily accomplished with the Multiplier and Electric Accounting Machine. The details of the machine process are fully described in Dr. Eckert's book *Punched Card Methods in Scientific Computation.*

The Electric Accounting Machine can be used to difference a punched-card table, directly computing and printing both first and second differences in a single, high-speed operation—2,400 per hour.

One use of punched-card tables is for the automatic application of values of functions to problem cards. This is done by sorting the problem cards in order according to the argument of the table and then automatically selecting the proper table cards with the Collator. The Reproducer then punches the data from the table onto the problem cards. If interpolation is required it is done by the Multiplier.

The table-making process is useful in statistics for converting raw scores. After the mean and standard deviation have been determined a linear equation of the form $Y=AX+B$ can be established and a punched-card table made for this function. The punched-card table is then sorted ahead of the raw score cards and the converted score gang punched into all cards.

Harmonic Functions. Several of the most important uses of the machines have been in connection with the synthesis and analysis of harmonic functions. A technical paper on the use of Hollerith machines for synthesis of harmonic series appeared in 1932. This paper, appearing in the *Monthly Notices of the Royal Astronomical Society,* presented by Mr. L. J. Comrie, described the method by which the many coefficients in Brown's tables of the moon were combined into an orbit for the

moon carried out to the year 2,000. This was the most difficult case of harmonic synthesis where the periods of the various components were not commensurable. This paper explains the method by which punched-card tables were prepared which took this fact into account and permitted the synthesis to proceed on a mechanical basis.

The use of cards for harmonic synthesis is indicated by the recurrent use of the same component values in different arrangements as the periods repeat. The preliminary card preparation consists in determining the interval of the desired synthesis and then computing the values of each term throughout its period for values of the argument at the desired interval. These values can be computed initially with the aid of a punched-card table and the multiplying punch.

The cards for each term are then placed in stacks on a table and the top card of each stack picked up and totaled in the Electric Accounting Machine. Checks on the proper selection of cards can be secured by adding card numbers as well as the coefficients. This addition can be performed automatically by placing a special card ahead of each group of cards for a distinct argument.

After tabulation, the cards for the various terms are separated by a run through the Sorter and replaced behind the unused cards for each respective term.

For repeated synthesis of Fourier Series, prepunched decks which can be combined to produce any amplitude of each frequency are used. Such a deck has been used extensively at California Institute of Technology and computes points at intervals of 1/500 of a circle and goes up to a frequency of 30.

In the analysis of harmonic series as distinct from synthesis, the Multiplier and Tabulator combine to give a most effective method in reducing the manual effort involved. These analyses are made with the aid of a set of prepunched cards. These cards are prepunched with the value of the trigonometric function and a pattern showing in successive columns whether the product formed on that card is to be added, subtracted or eliminated in computing the successive amplitude coefficients.

One interesting possibility in the use of cards for harmonic synthesis which was suggested by Comrie is that a whole series of harmonic syntheses involving the same periods but varying amplitudes can be tabulated from a double set of cards. By combining the original and duplicate sets of cards out of phase sufficiently, any desired amplitude from zero to twice the amplitude of the set may be produced.

Progressive Digiting. Tabulating machines have been recognized quite

widely as an efficient means of computing the sums of products needed in computation of multiple correlations and least square trend lines and other statistical and computational problems. This work can be performed with a sorter and the Electric Accounting Machine. The process is such that a number of cross products may be handled simultaneously in separate counter groups of the Electric Accounting Machine.

All the factors which are to be multiplied together are punched on cards, each card carrying the related data of a single case. The sums of the squares and of the cross products are obtained by a method of multiplication by addition. This process handles one multiplier digit at a time and is extremely rapid. Comrie states that "on one occasion 25,000 products of three-figure numbers were formed and added in about three hours." This method of multiplication is probably the fastest known today. Multiplication takes place at the same speed as addition and many products may be accumulated at one time.

A development of this method of multiplying by addition is now in use where the multiplication is done without sorting. In this process an analyzing device on the Electric Accounting Machine analyzes the digits of the multiplier column, adding or subtracting the multiplicand into one or more of three counters assigned values of 1, 3 and 5. For example, the digit 4 adds into counters 1 and 3, the digit 6 into counters 1 and 5, and digit 7 adds into counters 3 and 5 and is subtracted in counter 1, etc. These totals are summary punched and the 5 counter multiplied by 5; 3 counter by 3, and sum of 5, 3, 1 cross footed on multiplier.

Evaluation of Determinants. The solution of determinants, particularly of the higher orders, is one which is particularly burdensome when done manually. It is a problem to which the punched-card method has been applied for elimination of the manual labor. The method involves the use of the Multiplier, the Reproducer, and Sorter, and parallels the short method of single division usually followed under manual methods.

A card is punched for each element of the determinant identified by its row and column. The reciprocal of the element a_{11} is punched in card a_{11} and used as a group multiplier for all the cards in row 1. The cards for the remaining rows are offset gang punched transferring the value of the element of column 1 to the remaining cards for each row. The cards for row 1 are then sorted as group multiplier cards ahead of the remaining cards by column. The cards are then group multiplied for the reduction $a - b \times c$. This routine has reduced the determinant by one order and is repeated until the determinant can be evaluated at sight.

This method is entirely general, very rapid, and involves a limited number of simple machine operations. The identical process can be used for the solution of simultaneous equations.

Thus far we have discussed basic principles of punched-card methods for performing mathematical computations. In addition to the specific purposes of these basic principles there is a wide variety of problems which can be solved by combinations or repeated application of the basic punched-card steps. I will but briefly point out a few such problems which have been successfully attacked and solved in this manner.

One of the outstanding applications of the punched-card method is its application to the solution of differential equations which was developed by Dr. W. J. Eckert at Columbia. This machine procedure is extremely effective and we expect to see its application carried into many fields. Of particular significance in connection with this use of machines as compared with other methods is the degree of accuracy which can be established.

A simple but useful application of the Electric Accounting Machine is the preparation of *scatter diagrams*. The *punched-card technique* for factor analysis has been worked and successfully applied in Chicago.

Other major machine applications have been in the evaluation of formulae, for instance, the *transformation* of spherical coordinates into rectangular coordinates.

Another very extensive computation now being performed is a bivariate linear interpolation where a series of multiplier operations determines the weight to be given each of the four surrounding known values of the function and performs the final evaluation.

It should be clear that punched-card methods may be applied to many computational problems extensive enough to warrant mechanization. The three fundamental mathematical operations into which almost all computational problems can be transformed, namely, evaluation of determinants, evaluation of harmonic series, and evaluation of polynomials can be performed by these methods. Much of the preliminary work in applying punched-card methods to scientific computation has already been done by pioneers in this field. The task now before us is to exploit intensively the new methods and more efficient tools they have tested for us.

A COMPUTATIONAL SHORT CUT FOR REGRESSIONS BASED ON UNEQUAL FREQUENCIES[1]

By Marion M. Sandomire
Office of the Inspector-General, War Department

The regression of a set of quantities with unequal frequencies on corresponding equally spaced values of the independent variable may be computed by the ingenious method given in R. A. Fisher's *Statistical Methods for Research Workers*, Ed. 7, pp. 168–176. This method, in that it requires a number of successive additions, involves a great amount of writing. While it is true, as stated on page 170, "that much labour is saved by choosing a 'working zero'," it is of interest to eliminate all manual labor possible.

Since a method making use of orthogonal polynomials is to be preferred to one requiring cumulative frequencies, it is assumed that orthogonal polynomials are used in problems with equal frequencies. With unequal frequencies, short series with low degree regressions might possibly be written out without too much effort. With a large number of very long sets of observations, punch card equipment might be employed for obtaining the cumulative frequencies. However, without this equipment a calculating machine, especially one with multiplying keys, may be put to use most advantageously.

The following table gives factors that may be used with the values of the dependent variable and the frequencies to obtain successive cumulative sums without the intermediate recording that is shown in Fisher's tables 30.3 and 30.4. The method of constructing the table is apparent from the numbers themselves. A formal illustration with a simplified case as, for example, with four quantities above the working zero and four below follows. By writing out the successive additions above and below the zero line, quantities are obtained which are seen to be sums of products of the original values and the factors shown in the table.

	0	1	2
d	d	d	d
c	$c+d$	$c+2d$	$c+3d$
b	$b+c+d$	$b+2c+3d$	$b+3c+6d$
a	$a+b+c+d$	$a+2b+3c+4d$	$a+3b+6c+10d$
A	$A+B+C+D$		
B	$B+C+D$	$B+2C+3D$	
C	$C+D$	$C+2D$	$C+3D$
D	D	D	D

[1] The procedure described here was designed and applied when the writer was connected with the U. S. Forest Service.

MULTIPLIERS FOR OBTAINING CUMULATIVE SUMS

0	1	2	3	4	5	6
1	17	153	969	4845	20349	74613
1	16	136	816	3876	15504	54264
1	15	120	680	3060	11628	38760
1	14	105	560	2380	8568	27132
1	13	91	455	1820	6188	18564
1	12	78	364	1365	4368	12376
1	11	66	286	1001	3003	8008
1	10	55	220	715	2002	5005
1	9	45	165	495	1287	3003
1	8	36	120	330	792	1716
1	7	28	84	210	462	924
1	6	21	56	126	252	462
1	5	15	35	70	126	210
1	4	10	20	35	56	84
1	3	6	10	15	21	28
1	2	3	4	5	6	7
1	1	1	1	1	1	1
1						
1	1					
1	2	1				
1	3	3	1			
1	4	6	4	1		
1	5	10	10	5	1	
1	6	15	20	15	6	1
1	7	21	35	35	21	7
1	8	28	56	70	56	28
1	9	36	84	126	126	84
1	10	45	120	210	252	210
1	11	55	165	330	462	462
1	12	66	220	495	792	924
1	13	78	286	715	1287	1716
1	14	91	364	1001	2002	3003
1	15	105	455	1365	3003	5005
1	16	120	560	1820	4368	8008

The appearance of these factors from this arithmetic process formed the basis for proof, originally,[2] for this summation method of obtaining requisite products involving powers of the independent variable. This example is sufficiently general to indicate the method of constructing the entire table (which might be regarded as cumulations of unit frequencies).

The final quantities obtained in the forward and backward summing may be recorded, and then their sums and differences computed as indicated by Fisher (p. 171). However, it is readily seen that these results may be obtained directly in the calculating machine by taking the sign

[2] Proposed by G. F. Hardy and described by W. P. Elderton in Chap. 3 of his *Frequency Curves and Correlation*, 1938.

of the factors as negative in the odd columns above the zero line, and recording only the final result for each column. Using Fisher's data, for example, we obtain:

$$
\begin{array}{llr}
1{\cdot}6+1{\cdot}10+\ \cdot\ \cdot\ \cdot\ +1{\cdot}71+1{\cdot}67 & +1{\cdot}66+1{\cdot}61+\cdots+\ 1{\cdot}8+\ 1{\cdot}0= & 088 \\
0{\cdot}6+8{\cdot}10+\ \cdot\ \cdot\ \cdot\ +1{\cdot}71 & -1{\cdot}66-2{\cdot}51-\cdots-\ 0{\cdot}8-\ 10{\cdot}0= & 001 \\
30{\cdot}6+28{\cdot}10+\ \cdot\ \cdot\ +1{\cdot}66 & +1{\cdot}66+3{\cdot}61+\cdots+\ 46{\cdot}8+\ 66{\cdot}0= & 0054 \\
81{\cdot}6+60{\cdot}10+\cdots+1{\cdot}81 & -1{\cdot}66-4{\cdot}61-\cdots-165{\cdot}8-220{\cdot}0= & -3010
\end{array}
$$

and so on, for the remaining quantities.

If the factors are used as multipliers and accumulated, a check on their sum is obtained in the upper part of the table from the value found in the next column on the same line as the last factor used. In the lower part of the table, reference may be made to the above-mentioned value or to the value in the next column and one line below.

The table is written out in full to give the proper alignment as the columns drop down, and to have the factors arranged in the same order as the values themselves usually appear. The table is very easily extended, as needed, for either an increased range of the independent variable or for a higher degree curve.

The greatest advantage is obtained from this table of factors by transferring it to paper ruled in the same way as that on which the lists of values and frequencies are recorded. Alignment of the two permits the accumulation of cross products with little effort.

OBTAINING DIFFERENCES FROM PUNCHED CARDS

By Harry Pelle Hartkemeier and Herman E. Miller
University of Missouri

When statistical tables are available in the form of punched cards, it is frequently desirable to obtain first differences from such a table or set of punched cards. One method by which first differences can be obtained has been presented by W. J. Eckert in *Punched Card Methods in Scientific Computation.*[1] However, this method involves running the cards through the numerical tabulating machine twice because cards are tabulated in pairs, the first card in each pair being subtracted from the second card. An X must be punched in alternate cards. On the second run of the cards, the first card is removed in order to pair the second card with the third card, etc. A change must be made in the wiring of the X-distributor on the second run. Special wiring was used by Professor Eckert to obtain a break in control after every other card.

It is possible to obtain first differences on one run of the cards without any special control wiring when Type 405, the Alphabetic Accounting Machine[2] is used. Chart I shows the wiring necessary, and parts of the report appear in Chart II. Card columns 3, 4, 5, and 6 contain the argument which in this case is x/σ from 0 to 6.000. Card columns 7 through 13 contain the function,[3] which is the area under the normal curve between the arithmetic mean and the value of x/σ.

The value of the function punched in each card is added in one counter and subtracted from another counter. The value of the function punched in the next card is registered in the same two counters with the sign reversed by using the X punched in alternate cards to control an X-distributor. In this particular set of cards the X's were punched above the even numbers of the digit in column 6. All cards with odd numbers in column 6 are *NO-X* cards.

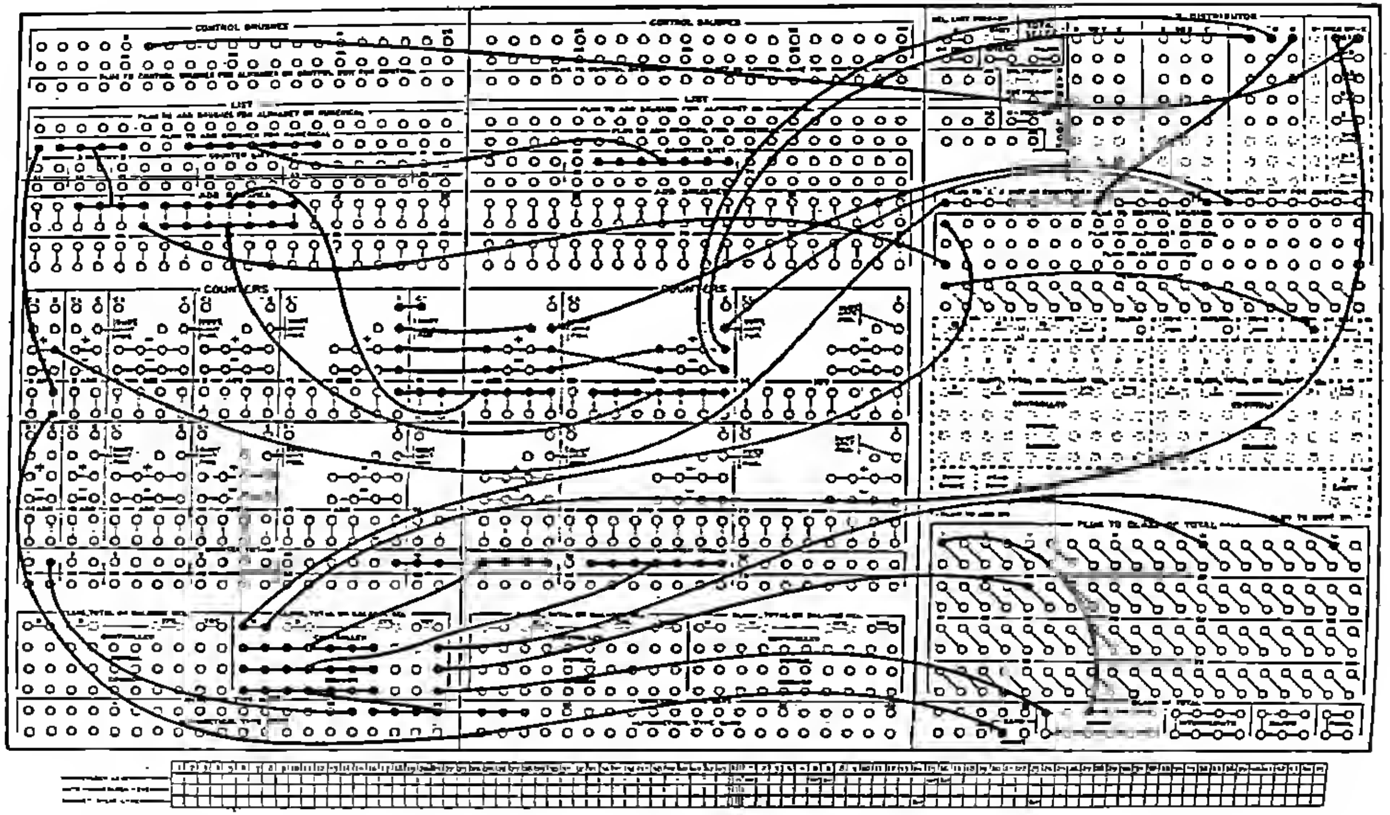
The X's punched in column 6 are also used to obtain control breaks after each card and to control the selector which prints totals alternately from the two counters. The class of total impulse is wired through the total selector so that the counters will be cleared alternately at the proper time. This wiring illustrates the great value of controlling total printing or counter clearance by wiring rather than by switches, as is done on Types 285 and 297.

[1] This book was reviewed in this Journal, Volume 36, Number 214, June, 1941, pp. 314–315.

[2] For detailed instructions and illustrations of how to operate this machine, see *Principles of Punch-Card Machine Operation*, by Harry Pelle Hartkemeier, New York, Thomas Y. Crowell Co., 1942.

[3] Quintic interpolation was used to compute values of the function in between those given in Karl Pearson's *Tables for Statisticians and Biometricians.*

CHART I—WIRING NECESSARY TO OBTAIN FIRST DIFFERENCES



The card count connection to type bar number 1 for listing numerical data causes zeros to be printed by type bars 2, 3, and 4. The card count connection to a counter enables us to register something in number 16 numerical type bar for total printing. We prevent the type bars actuated by the card count mechanism from printing by using hammerlock levers 1 and 16.

If second differences are desired, the first differences can be summary punched and used to obtain second differences. A check upon the punching can be obtained by noticing that the sum of the first differences approaches .5000000 as a limit. Counter coupling is used in Chart I because some Type 405 machines have only the first 32 counters wired.

CHART II

PARTS OF THE PRINTED REPORT OBTAINED FROM THE ALPHABETIC ACCOUNTING MACHINE

$\frac{x}{\sigma}$	Area	First Difference
0.000	.0000000	.0000000
0.001	.0003989	.0003989
0.002	.0007979	.0003990
0.003	.0011968	.0003989
0.004	.0015958	.0003990
0.005	.0019947	.0003989
0.006	.0023937	.0003990
0.007	.0027926	.0003989
0.008	.0031915	.0003989
0.009	.0035905	.0003990
0.010	.0039894	.0003989
0.011	.0043883	.0003989
0.012	.0047872	.0003989
0.013	.0051861	.0003989
.....	....	....
.....	....	....
.....	....	....
0.737	.2694388	.0003041
0.738	.2697428	.0003040
0.739	.2700465	.0003037
0.740	.2703500	.0003035
0.741	.2706533	.0003033
0.742	.2709563	.0003030
0.743	.2712591	.0003028
0.744	.2715617	.0003026
0.745	.2718641	.0003024
0.746	.2721663	.0003022
0.747	.2724682	.0003019
0.748	.2727699	.0003017
0.749	.2730714	.0003015
0.750	.2733726	.0003012

BOOK REVIEWS

GLENN E. MCLAUGHLIN
Review Editor

Agricultural Price Analysis, by Geoffrey S. Shepherd. Ames, Iowa: The Iowa State College Press. 1941. viii, 402 pp. $3.75.

For twenty years agricultural economists have been building up a technique of quantitative measurements of commodity supply-demand relations, almost unknown to economic workers in other fields. These studies have served as technical foundations and guides for vast public action programs in the field of farm production adjustment, price control, and agricultural planning. The number of published "price analysis" studies runs into the high hundreds, if not into the thousands. The books in this field have been few and far between, with Thomsen's elementary *Agricultural Prices*[1] and Henry Schultz's monumental *Theory and Measurement of Demand*[2] the major printed texts to record the generation of development that followed the first explorations of Henry L. Moore,[3] a generation ago.

Shepherd presents a treatment of the subject that is sophisticated from the economist's point of view. It ties back firmly to the institutional facts of market organization and to modern economic theory. It carries on from the price analyses themselves to their meaning in analyzing social action programs (such as the A.A.A., the Food Stamp Plan, milk price plans, and marketing agreements, and pro-rates) and even further to a consideration of the basic problem of how full employment can be maintained in a society where monopolistic competition is the rule and effective competition the exception.

After the excellent discussion of the market institutions within which farm prices are recorded, Shepherd gives a step by step discussion of the basic economic phases of price analysis. In this section his approach in terms of supply and demand curves and of their elasticities leads perhaps to an over-emphasis on those economic abstractions rather than to analysis of the fundamental problems of explaining changes in prices, changes in production, and changes in supply offered. The novice might wish also here for some clearer hints and instructions for actual research operations. For example, the chapter on elasticity of supply gives no clear statement of the fact that usually only by taking *changes* in production, rather than absolute amounts, has the agricultural supply curve been found measurable for successive production periods. Subsequent sections however, are quite practical. Statisticians will find of special interest the clear discussions of the use of multiple correlation and the graphic short-cut method in price analysis in Chapter 18, and of the significance of price analysis results in Chapter 21.

This is a useful and stimulating book, and one that may be an eye-opener

[1] F. L. Thomsen, *Agricultural Prices*, McGraw-Hill, 1936.
[2] Henry Schultz, *Theory and Measurement of Demand*, University of Chicago Press, 1938.
[3] Henry L. Moore, *Economic Cycles, Their Law and Cause*, Macmillan, 1914.

to economists in other fields who wonder how economic functional relations can be measured in quantitative terms. It summarizes an area of work that is moving continuously, though fumblingly at times, to provide economics with some of the same quantitative bases that the natural sciences have.

Mordecai Ezekiel

U. S. Department of Agriculture

A Theoretical Analysis of Imperfect Competition with Special Application to the Agricultural Industries, by William H. Nicholls. Ames, Iowa: The Iowa State College Press. 1941. xiv, 384 pp. $3.75.

Dr. Nicholls undertakes in this book to present the theory of monopolistic or imperfect competition in an agricultural setting. It is essentially a textbook for use in graduate classes in marketing and prices; that is, it is not a monograph in the sense of being an exhaustive inductive study of a particular situation. As a theoretical treatise it is not so much an attempt to carry further the Chamberlin-Robinson type of analysis as to restate it and, as Nicholls puts it, to assume the role of the tool-adapter.

The approach used is one obviously needed in graduate instruction for students of agricultural marketing. Many researchers will find very helpful the carefully developed theoretical framework which the book presents. Marketing textbooks very generally have stressed the institutional organization of the markets and the functions performed. Empirical researches more often than not have dealt with agricultural marketing as a production process rather than as a bargaining procedure. Prices tend to be taken as *faits accomplis* related to the supplies offered by original producers and the demands arising from ultimate consumers. Emphasis is thus on basic supply and ultimate demand, and upon efficiency in performance of the physical functions of marketing. In this connection, consideration is likely to be given to scale of operations in the handling agency as it affects the cost of processing and handling. Dr. Nicholls places emphasis upon the price-making aspect of marketing and possibly in some measure loses sight of social gains as well as losses which may arise from large-scale market operations. Such variations in cost of operation are, to be sure, included in his graphs but in exceedingly schematic form. They are not much discussed.

The fundamental interest of producers is to gain as large a portion as possible of what ultimate consumers can be induced to pay; of consumers to get their product at as little as possible above the amounts needed to induce production. Large-scale operations probably produce larger economies than most schematic presentations indicate. Granting, then, a tendency for this larger scale operation to open the way to certain monopolistic practices, the fundamental problems become those of appraising the net effects of few firms and large-scale operations and of ascertaining what measure of competition between them can and should be brought about, and how it is to be done. Nicholls recognizes this problem in a brief comment in Chapter 9 but

sets it aside rather quickly. He has, probably wisely, stuck to his main task, though this larger over-all problem will eventually need fuller exploration than it has had thus far.

Space available for this review does not permit extended comment on the carefully developed analysis covered. The book presents a new version of the Chamberlin-Robinson approach, and carries forward the attempt to apply the abstract theory developed by these earlier writers. The examples taken are more illustrative of what might be the competitive situation, or lack of it, than proof of what is. Many instructors and researchers, however, will find it a very valuable addition to their teaching and research materials.

Not a few students of agricultural marketing problems will no doubt object that the refinements described run much beyond the actual knowledge available even to monopolists and oligopolists, and that the many rigid assumptions necessitated by this type of analysis remove it from reality to an extent that vitiates its usefulness. Nevertheless Dr. Nicholls has made real progress in the direction of adequate analysis of these complex problems. It is a scholarly and competent exposition.

MURRAY R. BENEDICT

University of California, Berkeley

Wheat Studies of the Food Research Institute, Stanford University, California. "Wheat in National Diets," by M. K. Bennett. Volume XVIII, No. 2. October 1941. pp. 37–70. $1.00.

On 40 large pages the author succeeded to cover very much ground. Not only did he give a comprehensive picture of the role of wheat in the diet of as many as 52 countries, but in addition there is a computation of the cereal-potato ratio to total food consumption in calories. The reviewer found this supplement of especially great interest.

Bennett's computations necessarily had to be based on data of varying degrees of exactness, representativeness, and comparability. He had, for example, to use data on disappearance rather than actual consumption, i.e., his data include waste in trade channels and homes. It is common practice to make the necessary qualifications in footnotes which many do not care to read and which are mostly disregarded in reproductions of the data. You cannot do this with Bennett's data. His less reliable figures on total food disappearance in calories and on the cereal-potato ratio to total food are presented either with intervals of 200 calories per day for total food disappearance or in percentage ranges such as 50–60, 60–70, and the like for the cereal-potato ratios. These intervals and ranges are inserted in both the tables and charts and no one can fail to neglect them.

Bennett's data support by broad statistical analysis the commonly known fact that with due allowance for differences in average weight of the people and in climate, the variations in total intake of calories in the form of food are relatively small. Except for several Asiatic countries, total disappearance

of food is within the rather narrow range of 3,000 to 4,000 calories per adult male. However, the proportion of cereals and potatoes in total calories disappeared varies from 30–40 in the United States, Canada, United Kingdom, Sweden, Switzerland, Australia and New Zealand to as much as 80–90 in Russia, Rumania, several Asiatic countries, including China and India, and also in a few African countries.

The proportion of wheat in total calories disappeared naturally varies even much more than that of calories in cereals and potatoes, from almost half of all calories in Bulgaria to nothing in Nigeria and practically nothing in Java, French Indo-China, and several other countries.

The last section of the study compares the changes in per capita disappearance of wheat flour from 1923–28 to 1933–38 in the various countries. More countries show declines than rises and all the most important wheat consuming countries, such as the United States, United Kingdom, France, Italy, and Germany, are among those showing declines. (No such comparison could be made for Russia.)

The only doubt the reviewer has is the wisdom of computing the ratio of per-capita wheat flour disappearance to per-capita disappearance of total food in calories in Chart 2, by using not only the caloric value of actually disappeared total food but also a constant figure of 3,000 calories.

N. Jasny

Washington, D. C.

Statistical Methods for Research Workers (Eighth Edition), by R. A. Fisher. Edinburgh: Oliver and Boyd. 1941. xv, 344 pp. 10 shillings.

In preparing the eighth edition of this standard work, Professor Fisher has followed his usual custom of adding new sections describing more recently developed techniques, while making little or no change in the material contained in previous editions. The principal addition consists of an extension of the section on the use of discriminant functions. A discriminant function is that linear compound of a set of measurements which best distinguishes between a number of groups, in the sense that the ratio of the mean square between groups to the mean square within groups is maximized. In the seventh edition Fisher outlined the calculations required to construct the discriminant function, and described a number of practical applications of this new tool. The present edition contains a numerical example illustrating an approximate test of significance of the difference between the "best" discriminant and the discriminant obtained from any given linear compound of the measurements. Thus we may test whether the "best" discriminant differs significantly from an index which might be constructed either from theoretical considerations or by assigning on inspection an arbitrary set of weights to the various measurements. This test should be illuminating to those readers who wish to understand the relation between the discriminant function approach and previous methods of handling the same type of problem.

The preface contains an interesting discussion of the extent to which the book is dated. In the author's opinion, this is most evident in the order of presentation of the topics; in particular, he suggests that the analysis of variance might profitably be presented sooner and developed more fully. As a step in this direction, paragraphs have been inserted in this edition indicating the relation of the t-test to the analysis of variance tests of significance, though the detailed discussion of the analysis of variance still follows the section on intra-class correlations, and Professor Fisher excuses himself from the heavy task of a complete revision. Owing to the author's insistence on giving exact formulae rather than approximations wherever possible, the book remains as up-to-date a laboratory manual as any in the literature. From the point of view of its use as a text-book, however, the arguments in favor of a rearrangement will increase in force as the editions grow with the years.

Presumably on account of wartime restrictions on the use of paper, the type has been reset so that the lines of print are closer together. The size of type is unchanged, however, and the book is still very comfortable to the eye. Some printing errors have unfortunately crept in during the resetting; mistakes were noted in the formula for k_1 on p. 70, for a on p. 145, for z on p. 191, for the value given to the formal variate y in the group of males on p. 280, and for the additional information on p. 320.

W. G. COCHRAN

Iowa State College

Factor Analysis, by Karl J. Holzinger and Harry H. Harman. Chicago: University of Chicago Press. 1941. xii, 417 pp. $5.00.

Although factor analysis had its greatest development in the field of psychology, its application to many other branches of the social and physical sciences is becoming so wide-spread that a book dealing with the general aspects of this subject is not only timely but highly welcome.

The book under review is a valuable contribution for several other reasons: First, the authors have made a determined attempt to define the scope and limitations of the technique of factor analysis; second, they have synthesized the major known methods of factorial decomposition; and third, they have developed the subject matter with great clarity and with much detail so that the book can be followed without too much difficulty by people with limited mathematical training.

Except for a few minor omissions, the book is self-contained. Not only does it cover all of the important aspects of factor analysis that have been considered in the past few decades in the literature, but each topic is thoroughly digested with detailed numerical examples, detailed methods of computation, and complete proofs given either in the body of the text or in an appendix.

From the point of view of the uninitiated, the mathematical and geometrical treatment in this book of the subject of factor analysis is the best

that has yet appeared. The main contribution of this book, however, lies not so much in the algebraic and geometrical treatment as in the handling of the subject matter as a tool of research.

As is well known and is frequently emphasized by the authors, there exists no unique method of resolving a set of statistical variables into factors. (From an algebraic point of view, this is another way of saying that there exists no unique method of reducing a quadratic form into a sum of squares.) It follows therefore that the criterion for what the authors called a "preferred" system of factors must be based on considerations other than mathematical.

From a strictly scientific point of view, the ideal preferred system of factors is one which is deduced from an *a priori* hypothesis concerning the characteristics which the variables under analysis are supposed to measure. Once a system is derived in this manner, statistical tools can then be developed to test the agreement between the hypothesis and the observed facts. An outstanding example of this type of factor analysis is given by Spearman's "Bi-Factor" theory of intelligence.

Generally speaking, however, the technique of factor analysis has heretofore been used more as a tool for discovering hypotheses than for testing hypotheses which have been arrived at on *a priori* grounds. Such an application of this technique is necessarily full of pitfalls. The authors of this book have given a great deal of attention to this and similar problems and have arrived at several rational methods of judging the system of factors which is to be preferred.

Insofar as factor analysis is a tool of statistical analysis, the problem of the significance of the results obtained from a sample of observations assumes great importance. Unfortunately the problems involved in the development of tests of significance are rather complicated, and, to date, only scanty progress has been made in this field. The book, for example, contains several approximations to the standard errors of the quantities involved and some indications of tests of significance are given. However, these are at best rough and cannot be relied upon always to give satisfactory answers. Thus, at the present stage of development, an investigator must depend to a large extent on his own judgment and a knowledge of the variables analyzed. The same holds true in regard to the problem of estimating communalities. These communalities are in a sense the cornerstone of the whole technique of factor analysis as developed in this book, and several methods are described for their estimation. If used with discrimination and mature judgment, these methods will lead to fruitful and sensible results. This fact is well exemplified by the manner in which the authors themselves have handled them. However in the final analysis, the methods proposed are mostly based on rule-of-the-thumb criteria and will remain as such until factor analysis passes from a tool for hit-and-miss experimentation to a tool for testing well-thought-out theories concerning the structure of the variables analyzed.

M. A. Girshick

U. S. Department of Agriculture

The Analysis of Economic Time Series, by Harold T. Davis. Bloomington, Indiana: The Principia Press, Inc. 1941. xiv, 620 pp. $5.00.

The difficult problem of analyzing and interpreting economic time series has received its greatest attention during the past decade. Professor Davis, therefore, has rendered a distinct service in setting forth in this volume the present status of research in this field. The book is concerned chiefly with developing and analyzing the components of economic time series such as trends and cycles. In this connection considerable attention is given to the technique of harmonic analysis and its application to economic time series. The measurement of the energy in the motion of time series, or the variance attributable to the components relative to the total variance of the series is a central feature of the analysis. While the treatment of these problems is presented in some detail and with mathematical precision, important aspects of the subject are inadequately covered. Problems of discovering and interpreting relationships among time series and the associated problems of significance play but a minor and unsatisfactory part in the analysis. It must be pointed out that to follow the details of this work requires a considerable knowledge of mathematics, particularly function theory. The treatment of the material from this standpoint is elegant and those familiar with this field should find the book stimulating and thought provoking.

The first chapter of the book presents a concise history of the problem and serves as a splendid introduction to the treatment given in detail in the chapters that follow. Two chapters are devoted to the theory of harmonic analysis and its application. A discussion is given of the analysis of trends which includes applications of the logistic curve to the study of population growth. Other subjects treated in detail are serial correlation, theory of random series, the degrees of freedom in economic time series, and the nature of income and wealth.

A major interest in these techniques lies in their utility in forecasting economic variables, a problem considered in the closing chapters of the book. The discussion at this point is inadequate since it throws no light on the problems that must be solved in forecasting economic time series. Variations in the movement of such series are usually dependent on many factors some of which are measurable and others non-measurable. Criteria for the selection of the factors, the form of relationship among the variables, and tests of the reliability of the forecasts are essential in predicting the future of economic variables. This phase of the problem is not touched upon in this volume. Despite these shortcomings, Professor Davis has made a distinct contribution in making available to the student a brilliant presentation of powerful techniques of analysis which promise much for the future development of economics as a science.

LOUIS J. PARADISO

U. S. Department of Commerce

A First Course in Statistics, by E. F. Lindquist. Boston: Houghton Mifflin Company. 1942. xiii, 242 pp. $2.50.

Study Manual for a First Course in Statistics, by E. F. Lindquist. Boston: Houghton Mifflin Company. 1942. 117 pp. $1.00.

This text, with its lack of exercises, impresses one as being too verbal and descriptive. This can be explained, however, in the light of the purposes as stated in the preface. The author states that he wishes to "stress as much as possible the uses and interpretations and minimize as much as possible the mathematical theory of statistics and the mechanics of computation."

While the aims are worthy, it is difficult to see how a sound appreciation of statistical procedures can be achieved or thoroughly critical attitudes can be developed without some training in the mathematical bases of the subject. Furthermore, one may fairly question the inclusiveness of these aims, since it is highly desirable that students gain some adequate facility in the use of all basic processes. The correct choice of procedures and the precise manipulation of techniques are certainly essential aspects of interpretation.

The treatment is unique for an elementary text and one which is timely. The emphasis on student activity in logical analysis, the stress placed on sampling theory and especially the techniques suitable for small samples, the use of statistical methods, and the systematic development of essential information, make for a valuable contribution which may help to reshape textbooks in this field and on this level. The text is also to be commended for its clear definition of the "integral measure," its presentation of definite steps in computational procedures, its recognition of the fact that relatively few of the distributions actually encountered are "normal," and its common-sense interpretation of reliability and the coefficient of correlation.

There is a tendency toward formalization in treating the determination of the number and size of intervals, a lack of presentation of operations suitable for varied types of data such as attributes and skewed distributions, and a failure to acquaint the student with variations in processes of computation. Some may also question the advisability of omitting such topics as the harmonic mean, the geometric mean, binomial expansion, index numbers, non-linear correlation, and many of the applications of the normal curve theory used in test construction.

The manual, as indicated by the title and a statement in the Foreword, is designed for use with the text. One may doubt that it "cannot be used with other texts," as most of it should be clearly comprehensible to the careful student of other texts in education and psychology. One normally anticipates that a manual which accompanies a statistical text will be characterized by many specific exercises calling for much practice in computation. In this case, however, the manual is almost as verbal as the text, since the exercises, for the most part, call for verbal explanation and interpretation.

In places one detects a tendency for these questions to become highly involved in minor issues, or to deal with points which have little general

significance. The instructor, of course, has freedom of selection. The pages are perforated and punched, making it easy to hand the assignments in and later to bind them in a notebook.

Many experienced instructors find it difficult to adapt themselves to the use of a manual constructed by another, chiefly because of variance of convictions as to where emphases should be placed. The manual will prove of great assistance to those who wish to stress logical analysis and interpretation, even though they find it advisable to supplement it with practical exercises for the purpose of developing facility in operations by the student.

Paul V. West

New York University

Investment and Business Cycles, by James W. Angell. New York: McGraw-Hill. 1941. xviii, 363 pp. $3.00.

Professor Angell's interesting study falls into three logically distinct, but closely related parts: First he attempts to set up a general hypothesis which will explain the appearance and reappearance of what the author considers the core of observed cyclical fluctuations in modern economic activity, "self-generating business cycles"; then he proceeds to examine the relationships between changes in the money stock, the volume of new investment, national money income and the volume of employment, laying stress upon the circular velocity of money as an explanatory mechanism, but also giving an illuminating discussion of the "multiplier" concept; finally he discusses the problems connected with government spending and the financing of the defense program, as seen from the theoretical background given earlier.

Professor Angell rejects any attempt to explain the similar cyclical behavior found in many economic time series by reference to "exogenous factors," as was done by Schumpeter, Moore, et al. Instead he sees inherent in any system of relatively free enterprise a set of forces which make for persistent and roughly simultaneous up and down fluctuations in employment, national income, investment, etc., fluctuations which would occur in a never ending "self-generating" sequence even if all outside factors were stabilized. As he sees this mechanism, cyclical changes in national money income are mainly produced by changes in the volume of new investment; this latter depends in turn upon the general level of anticipations in the previous period; and this in turn hinges largely on the rate of change of income during a still earlier period. Considered from a mathematical point of view, these relationships do not necessarily imply a cyclical movement of income, but are also consistent with indefinite monotonic increase or decrease. The upward movement must come to an end because of some or all of three sets of factors: disproportionately rising cost-price ratios; saturation of short run investment activity because of accumulated investment in the recent past; and—presumably most fundamental—the fact that the middle and upper income classes cannot rationally continue indefinitely to spend

all their increases in income on consumption and investment, so that they must begin to hoard after a certain point. These all make for a decline in the rate of increase of incoms, a fall in the level of anticipations, and the start of the downturn. The converse process occurs in the transition from trough to upturn.

In the second part of this study, stress is laid upon the changed relationships between national income and money stock as between 1890–1920, 1929–33, and 1033–30. The altered money-using habits of the American people which these changes reflect, Professor Angell suggests somewhat tacitly, may be a fundamental clue to the altered nature of our economic problems since 1029.

The acute observations on the current scene, particularly the suggestion that systematic manipulation of government spending and taxing powers may be means of stabilizing economic activity, are extremely interesting, and offer valuable clues for possible government action to win the peace after the present conflict.

There is little one can say in just criticism which has not already been anticipated and considered by the author. He admits he does not deal with the whole of cyclical economic reality. In fact, in his later chapters Professor Angell lays to deficit spending main responsibility for the 1034–36 upturn in this country. He recognizes the difficulty of statistical verification of his hypothesis, and concedes that the numerical coefficients defining his model relationships would probably differ considerably from period to period, reflecting largely the impact of the exogenous forces.

As the United States goes through its present trial, and we see the all too patent relationship between the present upturn in production and employment, and a very specific "outside" factor, there is a temptation to reject off hand any such "self-generating" hypothesis as offered here. Moreover it seems unlikely that in the future there will be quite the same type of economic freedom as that which Professor Angell sees as having a primary causal significance in making the downturn inevitable. Nevertheless this study has value as an attempt toward logical explanation of the cyclical activity experienced under capitalism and perhaps as a guide to future action to modify the "self-generating" possibilities of an economy in which there is both inequality of income and freedom of investment opportunity.

Harry Schwartz

Brooklyn College and War Production Board

Deficit Spending and the National Income, by Henry H. Villard. New York: Farrar and Rinehart. 1041. xviii, 429 pp. $3.50.

The economic effects of public expenditure when resources are fully employed, as they tend to be in a wartime economy, are naturally quite different from its effects in time of depression. Mr. Villard's book represents the most comprehensive review of the latter problem, both theoretical and

practical, which has so far appeared. In periods of underemployment may not government spending prove an effective method of raising the level of economic activity, perhaps even of securing permanent prosperity? This question is scarcely topical at the moment, but its practical interest seems certain to be revived in the future, perhaps suddenly and acutely. After a review, both thorough and discriminating, of the analytical literature of the past decade relating to the multiplier, the author declares that the case for cyclical deficit spending has been established. This is defined as spending which leaves the national debt unchanged over the course of an entire business cycle. On the other hand he rejects the plea sometimes offered that continuous or progressively increasing public investment is necessary to offset the decline in population growth and other long term factors leading to the stagnation of investment, preferring if necessary to discourage the propensity to save by appropriate measures of taxation; "only to the extent that a reduction in . . . savings is unable to solve the problem of secular unemployment does there seem to be a case for permanent deficits." However, he would presumably admit improvements financed by loans which led to secular increases in the national debt if such improvements were desirable in themselves.

To the statistician the main interest of the book will probably center in Villard's derivation of monthly estimates of net income-increasing expenditure for all governmental units for the period since 1929. In the case of the federal component, the adjustments made to the crude cash deficit (to exclude revenue which is not income-decreasing and expenditure which is not income-increasing) are for the first time explained and justified in detail. The new monthly series for the combined cash deficit of state and local governments, derived from changes in outstanding debt, will be of great value even though the coverage of cash assets and short term liabilities is admittedly incomplete. The substantial discrepancies between Villard's series for this item (Table 15) and the official estimates on an annual basis presented to the Temporary National Economic Committee (original data, *Hearings*, p. 4011 (Part 9); revised data, *Monograph 37*, p. 111) are disturbing. However, the facts that the Treasury have not published their methods, and that Villard is able notwithstanding to make what appear to be damaging criticisms of these methods, suggest that his estimates are preferable to the official ones.

HAROLD BARGER

Columbia University

Paying for Defense, by A. G. Hart, E. D. Allen, and others. Philadelphia: The Blakiston Company. 1941. viii, 272 pp. $2.50.

Paying for Defense in substance does two things: it sets forth the immediate economic and financial problems which the war has presented to the people of the United States, and it indicates the direction in which the

authors believe the answers are to be found. The book is simply and directly written, and while much careful and scholarly research went into it, it is clearly designed for the layman as well as for the economist. As must be the case with a volume that deals with a rapidly changing current situation, certain of the facts and figures are somewhat out of date, but this circumstance does not diminish the usefulness of the study.

The presentation of the basic problem of war economics, how to furnish the war effort with the necessary men and material and at the same time provide the requisite funds in such a way as to minimize price increases and inflationary dangers, is excellently done. Few persons will cavil at the authors' definition of war objectives—getting maximum output; preventing inflationary general price increases; sharing defense burdens fairly; giving all citizens a sense of sharing in defense; releasing resources needed for defense; promoting a healthy financial structure. In analyzing the possible ways of achieving these goals the authors bring to bear their familiarity with recent economic thinking and research, and explore the different types of taxation and borrowing, indicating the chief advantages and disadvantages of each.

The authors believe that the country is most likely to obtain the desired objectives through a broadening of the tax base, a deduction at the source income tax, and a flexible income tax rate which would be adjusted upward when increases in the cost of living evidenced the presence of excess purchasing power and the desirability that it be absorbed by the government. The economic reasoning that underlies this predilection is hardly to be questioned. Yet it may be doubted whether the analysis explores sufficiently the administrative feasibility of the proposal, and whether the study gives enough attention to the difficulties that business concerns would face were they suddenly to find themselves the principal tax collectors of the Federal Government. These topics are considered in the book; but the very real problems which lie in these areas are not to be minimized.

The authors are to be congratulated both on the skill and the speed with which they have analyzed in terms that the general public can comprehend the issues of war economics.

Charles Cortez Abbott

Harvard University Graduate School of Business Administration

---

The Flow of Business Funds and Consumer Purchasing Power, by Ruth P. Mack. New York: Columbia University Press. 1941. xvii, 400 pp. $3.75.

In this long and rather difficult to read book, Mrs. Mack analyzes the relationship between the flow of funds through business enterprises and the volume of consumer purchasing power. Chapter I is introductory. Chapters II–VI, occupying half the book, discuss selected sources and uses of funds and changes in some balance sheet items of fifty-four corporations, princi-

pally for the years 1934–38. This analysis is unusually full and detailed. One of the principal findings, that gross retained income during 1934–38 was more than adequate to meet total capital expenditures related to that income—despite the accumulated replacement need, due to wear and tear during the depression, despite rapid technological changes, and despite the not unusually high profits for the period—supports a similar proposition presented in the Temporary National Economic Committee's Hearings on Savings and Investment. Chapters VII and VIII, occupying one-quarter of the book, are based largely on interviews with eighty-six corporation executives. These chapters get at the "judgment factors and business techniques determining the depreciation accrual and capitalized expenditures for fixed assets"; they are extremely rich and interesting. Chapters IX and X, occupying the last quarter of the book, attempt to sketch the theoretical significance of the flow-of-income analysis, and the possibilities of using this analysis for maintaining full employment. Though these chapters are somewhat jerky and unpolished, they are probably the most important in the book.

Mrs. Mack is principally concerned with investigating the hypothesis that non-financial business enterprises may feed purchasing power to consumers at different rates than they are prepared to offer consumers' goods for sale. She is also concerned with the larger question whether consumer purchasing power in general is adequate. To analyze these propositions, Mrs. Mack develops the conception of "markets," or "economic areas within which income actually constitutes likely spending power in the same field." "The relations between income and expenditure within markets have certain characteristic patterns that may not be shared by other markets." Five markets are identified: consumer, industrial, financial, government, and foreign. The suggestion is made "that the transmarket spending of income is far more erratic and jumpy than is, for example, the flow of consumer income into consumers' purchases, or money invested in one sort of security into other securities or banks."

These thought-provoking hypotheses are not, however, conclusively tested. The calculations in Chapter IX, dealing with the estimated source or application of funds between the consumer and industrial market during 1932–38 are admittedly rough and based upon arbitrary allocations. On the other hand, the proposals for analyzing the flow of income with quarterly and monthly data are very suggestive.

OSCAR L. ALTMAN

National Resources Planning Board

Banking Operations in Ohio, 1920–1940, by J. M. Whitsett. Columbus: The Ohio State University, Bureau of Business Research. 1941. xxv, 217 pp.

This monograph presents a statistical study of the operations of state and national banks in the State of Ohio for the years 1920 to 1940. The statistical

tables presented are compiled from published reports of the Comptroller of the Currency and from both published and unpublished material in the Division of Banks of the State of Ohio.

There are six chapters dealing with (1) banks and the economic structure of Ohio; (2) changes in the banking structure of Ohio since 1920; (3) changes in banking practice in Ohio; (4) trends in earnings, expenses, and profits of Ohio state and national banks, 1921–1939; and (5) profitableness of Ohio state banks by size groups, 1930–1939.

This list of subjects represents an ambitious undertaking; it is not surprising perhaps that it raises hopes in the reader which are disappointed in the reading. The intriguing subject of "banks and the economic structure of Ohio," for example, is disappointingly discussed with some six pages of text in which for the most part materials are summarized from published sources that are a decade or more of age, and some of which represent broad generalizations on the national economy. The applicability of these generalizations to the economy of Ohio is not tested or commented on. The significance of "Ohio as an agricultural state" to banking is given only 16 lines, and "changes in marketing" (presumably in Ohio) are presented by a summary of the conclusions in *Recent Economic Changes*, published in 1929!

The chapters on banking operations and practices present material on number and resources of banks, the character of their assets and liabilities, and their earnings, expenses, and operating ratios, broken down by state and national bank classes; this material should provide both interesting and useful standards by which individual banks in the state could appraise their own position and performance. The treatment of these materials in the text is somewhat cumbersome, but there are a number of skillfully drawn charts which facilitate interpretation.

Ernest M. Fisher

American Bankers Association

Exchange Control and the Argentine Market, by Virgil Salera. New York: Columbia University Press. 1941. 283 pp. $3.50.

Shortly after the conclusion of the Ottawa Agreements a treaty was negotiated between Great Britain and Argentina providing for preferential treatment of Argentine meat exports and for a virtual balancing of the two countries' accounts which previously had been heavily in favor of Argentina. The main purpose of Salera's book is to show that the Argentine exchange control system was employed as a means of favoring British exports and discriminating against the United States by allotting favorable exchange rates to the former while forcing the latter to come in at the less favorable "free market" rate.

These discriminatory measures frequently appeared in disguise, and Salera does an excellent job in uncovering their effect upon the competitive status of Britsh and American products in the Argentine market. In general

the book provides much information that is becoming increasingly important to us at the present time.

The author's aversion to bilateralism, however, leads him to do less than justice to Argentina's economic policies. It is true, of course, that Argentina failed to obtain many of her imports from the cheapest supplier. On the other hand, the close association with the prosperous Sterling Block brought advantages which a trade agreement with the United States could never have matched. It is only since the outbreak of war that her dependence on Europe has become a serious drawback for Argentina.

In the handling of monetary and exchange problems the Argentine Government has shown ability of a very high order under difficult conditions. All through the depression Argentina was dependent on the export of agricultural products, while protectionism on the part of her customers increased. Nevertheless Argentine, by common consent, weathered the de-depression better than most other countries. She maintained a fairly stable price level throughout this period and managed to service her foreign debt without interruption. To some extent this performance, almost unmatched in Latin America, was due to fortunate circumstances, such as good harvests in years when there were crop failures in the northern hemisphere. But beyond this it has been due to three deliberate acts of policy: (1) abandoning the gold standard as early as 1929, thereby shielding the country against deflation, (2) the adoption of the dual market system of exchange control, which provided a safety valve in the form of the "free market" rate and prevented the creation of a rigid and artificial price system, and (3) the close association with the Sterling Block. A somewhat more appreciative evaluation of these policies would have added to the merits of Salera's book.

HENRY C. WALLICH

Federal Reserve Bank of New York

Interamerican Statistical Yearbook, 1940. Edited by Raul C. Migone with the assistance of Marcelo Aberastury, Emilio Fuento, and Jorge E. Iturraspe, under the auspices of the Comision Argentina De Altos Estudios Internacionales. New York: The Macmillan Company; Buenos Aires: El Ateneo; Rio de Janeiro: Freitas Bastos & Cia. 1940. 612 pp.

The *Interamerican Statistical Yearbook, 1940,* is a particularly useful pioneer venture undertaken under the auspices of the Argentine Commission of International Studies by a staff directed by the distinguished Argentine official and scholar, Dr. Raul C. Migone. Its announced purpose is to "serve the American community, holding before it a faithful mirror in which America may see herself as she is" (p. 24).

The data used were taken in most instances from publications of international organizations rather than directly from reports of any one country. The editors state that "all the data possible to obtain from official inter-

national sources" have been included in the volume. The content therefore reflects the relative availability of data on different subjects in such sources. About half of the volume consists of tables on international trade. The remaining half includes large sections on public finance, demography, and agricultural, mineral, and manufacturing production; and smaller sections on wages and hours of labor, prices, banking and money, transportation and communication, education, health, defense, and participation in international organizations and agreements. Subjects that are omitted because of lack of international data on a comparable basis include, among others, domestic trade, national income, occupations and employment, construction and housing, elections and extent of domestic political units, publications and libraries, and the extent of cultural organizations for concerts, theaters, and the fine arts. In arrangement, in attractiveness and presentation, the compilers are to be commended. The large clear print is also notable.

The textual matter, including heading, stub, and notes, is in each of the four official languages that exist in the American countries; namely, Spanish English, Portuguese, and French. While the problems of arrangement and space occasioned by use of four languages for purposes of international courtesy and increased popular usefulness have on the whole been admirably solved, the elimination of English and French in future editions should be considered. Advantages in added space for statistical content might outweigh the bother of dictionary reference by the few users who cannot read either Spanish or Portuguese, and the present text in British idiom is not always useful to North American readers.

There is a question whether the distinguished and competent editors might not discard advantageously their self-imposed handicap of depending almost wholly upon other international statistical compilations for their data, and undertake to obtain as much data as possible directly from the American governments or from their official reports. Although some additional work would be involved, the advantages of having a new independent source of reliable statistical information about the American nations would be enormous. Data would be more recent than is possible when obtained indirectly through other published compilations. Significant non-confidential data that exist only in relatively inaccessible or unpublished form in the capitals of the respective American countries would become available to the world at large for the first time. Perhaps also some nations might even begin the collection of information never before obtained by them, if requested to fill gaps in inter-American comparisons regularly included in an authoritative statistical yearbook dedicated to serve the American community. In any case, the *Yearbook* has become, with its first issue, a principal source of inter-American information collated with international perspective and high standards of scholarship.

E. R. Gray

Bureau of the Census

Commodity Yearbook, 1941. New York: Commodity Research Bureau, Inc. 1941. 936 pp. $7.50.

This reference work is the third in a series designed to provide both descriptive background and statistical records covering a wide range of commodities. It is prepared by a group of editors in New York who have had considerable experience in commodity markets, and who have utilized a variety of sources, official and unofficial, for this compilation. Each yearbook has given special emphasis to some aspect of commodity markets. Thus the 1940 edition covered up-to-date processing methods by which raw materials are converted into finished form, and the 1941 edition, in addition to a brief description of each commodity, discusses the source of supply of raw materials, the channels through which they are marketed, and some of the factors affecting supply and demand. The individual commodity sections, of which there are 75 in the 1941 *Yearbook*, vary in length from 3 or 4 pages for the less important commodities, to 10 to 25 pages for major markets such as wheat and flour, cotton, and steel. The statistical tables, which cover production, consumption, exports and imports, stocks, prices, and other relevant data, are exceedingly useful for reference purposes. The descriptive material, written in a highly readable style, is by no means exhaustive but indicates as a rule a considerable degree of familiarity on the part of the editor with the commodity in question and covers the high lights of the market. While such a yearbook cannot, of course, provide any great detail, it does give within the covers of a volume answers to a great many questions which sometimes require hours of searching through scattered sources.

The *Commodity Yearbook* for 1941 also has an excellent introductory section composed of three chapters. The first, on war-time control of commodities, summarizes the economic problems arising out of commodity regulation in World War I and the manner in which they were handled by various governmental agencies; the second is a clear description of the operation of commodity exchanges; and the third, on war and commodity prices, indicates the major factors affecting prices in war time.

Beyond these introductory chapters, there is little orientation toward the problems arising out of the war, since the *Yearbook* went to press early in 1941, and in this respect it may prove disappointing. Practically speaking, the statistical record ends with the year 1940, although some of the statistical tables cover the first quarter of 1941. In reading the individual commodity chapters in the light of the development of the defense program, there is a notable absence of a critical analysis of the shortages of supply which war abroad and defense at home have produced. This is especially true in the case of some of the metals—iron and steel, lead, aluminum. It is recognized, of course, that the industrial developments of 1940–41, after the beginning of the defense program, have greatly changed the supply picture. For a few commodities, moreover, certain phases of marketing and supply

appear not to have been adequately dealt with, although these are the exception rather than the rule. Milk is an illustration. Here neither the complex price-fixing procedure of state and federal regulatory agencies nor the buying policy of milk distributors is given adequate attention. From the point of view of the technical statistician, the tables would be improved by somewhat longer explanatory footnotes.

It is to be hoped that the 1942 edition will contain as much material as possible on the impact of the war on specific commodity markets, both as regards supply and demand and current techniques of government regulation. There is no doubt that the public will be greatly concerned with commodities for the duration of the war and that a current, up-to-date compendium of this kind should prove invaluable.

Aryness Joy

Bureau of Labor Statistics

Trends in Retail Trade and Consumer Buying Habits in the Metropolitan Boston Retail Area, by Richard P. Doherty. Boston: Bureau of Business Research, Boston University College of Business Administration, 1941. 40 pp. $1.00.

This study is an attempt to explain, on the basis of a consumer survey, the decentralization of retail sales away from the downtown stores in Greater Boston, as shown by the Censuses of Business for 1929, 1935, and 1939. The report offers the following substantive conclusions:

1. The Census Metropolitan District area is smaller than the area influenced by retail outlets in Boston proper.
2. Although the relative population of Boston and the communities included in a 30 mile radius have remained constant over the ten year period, there has been a decline in the percentage of sales accounted for by Boston stores.

The distribution of food sales has remained almost constant, but other classifications have shown an appreciable shift to stores outside of Boston proper, particularly in the case of automobile sales. The shifts in trade were presumably away from stores in the City of Boston to stores in the adjacent neighborhood area.

The consumer survey develops five reasons for consumer purchases in Boston stores by residents of outlying communities: better style, larger assortments, better quality, reliability of store, and price, including bargain sales.

Six reasons are given for purchases in stores within the "home" area: convenience in purchasing near home, reliability of stores, quality and style, delivery and charge service, acquaintance with proprietor or local loyalty, and location of stores with regard to theatre and other shopping facilities. A more detailed analysis of these shopping motives is given for each of the various commodity lines studied.

The report offers little new in technique, but, in its analysis of census data, offers a convenient basis for further analysis and application of information. Were similar reports made for other communities, the data would be even more useful. The biggest question in the mind of the reviewer has to do with the consumer survey which reinforces the analysis of census data. A questionnaire is mentioned, but no copy is submitted. No statement is made as to whether the survey is based on interviews or on a mail canvass. No statement is given as to the source of the list or the basis of distribution of interviews. If interviewers were used, the reader would be more confident if some word was given as to their selection, training, and supervision. Moreover, the consumer reactions reported are so general that some doubt may be justified as to whether the interviewing technique penetrated below rationalization.

LAWRENCE C. LOCKLEY

Curtis Publishing Company

Statistical Cost Functions of a Hosiery Mill, by Joel Dean. Chicago: The School of Business, University of Chicago Press. 1941. ix, 106 pp. $1.00.

This publication presents the results of one of a series of investigations by Professor Dean of the cost behavior of individual enterprises "with a view to obtaining valid generalizations concerning cost behavior under varying conditions." The study follows the pattern of multiple and partial correlation analyses heretofore used principally in demand function studies. Regressions were computed between output, and overhead cost, non-productive labor cost, productive labor cost, and combined total cost. Linearity was assumed in each case, and time was used as a third variable in each set of correlations. The partial regressions of time on the dependent or cost variable enabled the author to remove the influence of improvements in managerial performance and to reduce the scatter about the output-on-cost regression lines. Estimates of marginal and average costs were then made directly from the regression functions.

A careful series of tests was made for the reliability of the statistical analysis including the assumptions of linearity. Adjustments to remove the effect of changing input prices on costs had also to be provided where necessary.

Although not introducing any techniques basically new, this work presents the application of a simple statistical method to what is usually conceived a highly complex problem with apparently significant results. This end was of course facilitated by the simplicity of the production process considered. The elimination from costs of such items as depreciation, general administrative, and freight costs also tended to reduce the relative importance of overhead which otherwise might well have disturbed the linearity of the combined total cost-output regression.

O. J. MCDIARMID

College of William and Mary

The Petroleum Industry, by Ronald B. Shuman. Norman, Oklahoma: University of Oklahoma Press. 1941. xiv, 297 pp. $3.00.

Control of the Petroleum Industry by Major Oil Companies, by Roy C. Cook. Washington: Temporary National Economic Committee. Monograph No. 39. xi, 191 pp. $1.00.

Review and Criticism on Behalf of Standard Oil Company (New Jersey) and Sun Oil Company of Monograph No. 39, with Rejoinder by Monograph Author. Temporary National Economic Committee. Monograph No. 39-A. vi, 90 pp.

Professor Shuman's book is a brief survey of the oil industry: production, refining, transportation, marketing, taxation, labor problems, international trade in oil products, and conservation, with chapters on the demand for oil products and on the natural-gas industry. It is not exhaustive in its discussion of any of these subjects, but the author has given a clear and well-balanced picture of the industry, with judicious conclusions on controversial points. Professor Shuman offers no particular indictment of the Standard Oil Company, or of the twenty major oil companies, or of the oil industry in general, but sketches briefly the reasons why the integrated major companies dominate the industry, and why the independents live precariously and often not long. Perhaps it is ungrateful to wish that the book were larger and more complete, but the reviewer closed it wishing for more. The author knows the oil industry well, and he writes with clarity, pungency, and occasionally with fine humor. One of the amusing anecdotes in the book, for instance, relates to the corruption of state officials in oil-proration, when one firm was "credited with attempting to bribe a functionary to remain honest, the intent being to circumvent the plans of a rival to move out several times his allowable." *The Petroleum Industry* is a much more than ordinarily readable book.

Mr. Cook, a member of the staff of the Department of Justice, worked with the T.N.E.C. in the investigation of the oil industry, and his monograph is a study of the evidence presented before the T.N.E.C., and of other sources, a study devoted to the problem of monopoly control in the industry. He summarized the evidence of monopoly control by the integrated major companies, their 85 per cent of the trunk pipe line mileage, 96 per cent of the gasoline pipe line mileage, 87 per cent of the oil tankers, 75.6 per cent of the refining capacity, their almost complete control of patent companies, and their 52 per cent of crude production. He indicates furthermore, that their share of the business has been growing, and is likely to grow in the future. The marketing field, the most competitive part of the oil business, is left largely to individual operators, who are nevertheless under fairly complete control of the major companies, and apparently conduct their business at a loss. The author does not establish any *general* plan of cooperative or collusive action by the major companies, except in the activities of the American Petroleum Institute; but he does point out that they, or some of them, often

act cooperatively, as for instance, in the ownership of patents and pipe lines, and in some other matters.

The reply of U. S. Farish, of the Standard Oil Company, and J. Howard Pew, of the Sun Oil Company, is a general denial of the charge of monopoly, and of some of the conclusions of Mr. Cook; and Mr. Cook's rojoinder, insisting on the soundness of his analysis, is a part of the same monograph.

There is not space here to weigh all the evidence presented in these two monographs, but the reviewer would like to suggest that in their main contention, as to the question of monopoly, Mr. Cook and the oil men are both correct. Certainly the twenty major integrated companies do in a large measure "control" the oil industry; yet, just as certainly, there is competition, sometimes aggressive competition, among the majors and with the independents. In such competition the unintegrated independents are gravely handicapped, usually, and he must be an optimist who can see a bright future for them; but that situation is characteristic of many kinds of business today. The concept of imperfect competition would fit nicely here, or perhaps "monopolistic" competition would be better, with reasonable emphasis on monopolistic elements in the business. At any rate, these two monographs will be welcomed by students of the oil industry.

JOHN ISE

University of Kansas

Eleven Twenty-Six, A Decade of Social Science Research, by Louis Wirth, Editor. Chicago: University of Chicago Press. 1940. xv, 498 pp. $3.50.

This book depicts the story of ten years of activity at the Social Science Research Building of the University of Chicago—1126 East 59th Street. It contains the addresses and papers given at the tenth anniversary celebration (in December, 1939) of the dedication of the building, the discussions at five round tables held at the time, and a bibliography of the publications of present and past members of the University of Chicago social science faculty.

The volume is essentially a report of progress during a most important period in the development of social science in the United States. It shows the methods employed by outstanding social scientists, the wide range and interrelation of current problems, and the efforts being put forth to provide solutions.

More than half of the book is devoted to the proceedings of the tenth anniversary meeting, among which the round table discussions are the most significant, especially as they threw light upon insistent questions still in dispute among social scientists. Two such insistent questions, out of a number that are discussed in the proceedings, may be cited as illustrations.

(1) Should quantification in the social sciences proceed wholly from the point of view of ordinary "extensive" measurement or are there other quantification criteria of even greater importance here?

It was as a result of developments in modern physics, biology, and psy-

chology that the distinction between primary, secondary, and tertiary qualities or levels of analysis was advanced. From the point of view of quantitative analysis, the first level may be called the primary physical-science level of extension, form, and time. The second level may be designated the bio-psychological level dealing with degrees of difference within "subjectively" perceived qualities such as sound, color, taste, or smell. The third level may be called the psycho-social level covering relational differences among such qualities as desires, interests, judgments, and values.

(2) Should normative considerations have a place in social science or are they contra-indicated here as in natural science?

It is repeatedly insisted that we should not be concerned about doing "good" in the social studies. And yet it should also be borne in mind here that there is what seems to be an essential difference in this respect between the social studies and the natural sciences. In the social realm we are continuously doing "good" or "bad," not in our investigations necessarily, but as active agents who are constantly changing the social structure as such. We make governmental rules and regulations, for example. We pass laws, municipal, state and federal. And thus, in these and in many other respects, we are shaping and reshaping the social order itself. We make the social facts as we go, and we do not do that in the natural realm, that is, if we are thinking of the basic structure of the material universe, of the astronomical relations between the planets, the physical relations in mechanics, the interaction of the elements in chemistry, the constitution of the geological world, the genesis and evolution of the biological organisms with which we have contact. These are not changed by our being in the world in the sense that we have any power to change the laws governing them. But in being in the world (in historic time), man originally created the social order itself, and he has since repeatedly changed its essential constitution.

The question here is whether what we investigate in the social realm is not something much more flexible and unpredictable than what we investigate in the natural realm. In physical and biological science we can confidently predict what is to happen as soon as we have ascertained the laws governing a certain set of natural phenomena, such as Kepler discovered with respect to planetary motion. But in the social studies we have no such extensive power of prediction. At the same time, if man does have power to change the social order itself, then what is "good" or "bad" for society becomes a legitimate branch of social study. Ethics and normative standards may thus be significant for social science, whereas they may have no place in physical and biological science.

Eleven Twenty-Six contains discussions by able social scientists on many insistent current questions such as these. It will be welcomed by students and scholars interested in border-line problems and in keeping abreast of informed opinion regarding them.

Joseph Mayer

Bureau of Labor Statistics

PUBLICATIONS RECEIVED*

Allport, Gordon W. *The Use of Personal Documents in Psychological Science.* New York: Social Science Research Council. Bulletin 49. 1942. xix, 210 pp. $1.50

American Public Health Association, The Committee on the Hygiene of Housing of the. *Housing for Health.* Lancaster, Pennsylvania: The Science Press Printing Company. 1941. 221 pp. $1.00.

Bacon, Marvin A. *Income as an Index of the Fiscal Capacity of Michigan Counties.* Ann Arbor: University of Michigan Press. Michigan Governmental Studies, No. 8. 1941. viii, 78 pp.

Birkhoff, Garrett, and MacLane, Saunders. *A Survey of Modern Algebra.* New York: The Macmillan Company. 1941. xi, 450 pp. $3.75.

Blaser, Arthur Frederick, Jr. *The Federal Reserve Bank of Cleveland.* New York: Columbia University Press. 1942. xxvii, 309 pp. $3.50.

Boone, Gladys. *The Women's Trade Union Leagues in Great Britain and the United States of America.* New York: Columbia University Press. 1942. 283 pp. $3.50.

Borden, Neil H. *The Economic Effects of Advertising.* Chicago: Richard D. Irwin, Inc. 1942. xl, 988 pp. $5.00.

Chase, Stuart. *The Road We Are Traveling 1914–1942.* New York: The Twentieth Century Fund. 1942. 106 pp. $1.00.

Davis, Harold T. *The Analysis of Economic Time Series.* Bloomington, Indiana: The Principia Press, Inc. 1941. xiv, 620 pp. $5.00.

Davis, Hiram S. *Inventory Trends in Textile Production and Distribution.* Washington: The Textile Foundation. Number Seven of Inventory Policies in the Textile Industries. 1941. xv, 70 pp. 50 cents.

Day, John. *An Economic History of Athens Under Roman Domination.* New York: Columbia University Press. 1942. x, 300 pp. $3.50.

Dearing, Charles L. *American Highway Policy.* Washington: The Brookings Institution. 1941. xi, 286 pp. $3.00.

Dodd, Stuart Carter. *Dimensions of Society.* New York: The Macmillan Company. 1942. ix, 955 pp. $12.00.

Due, John F. *The Theory of Incidence of Sales Taxation.* New York: King's Crown Press. 1942. xii, 257 pp. $2.25.

Foulke, Roy A. *Credit Problems in a War Economy.* New York: Dun & Bradstreet, Inc. 1942. 71 pp.

Hartkemeier, Harry Pelle. *Principles of Punch-Card Machine Operation.* New York: Thomas Y. Crowell Company. 1942. xiv, 269 pp. $3.25.

Hultgren, Thor. *Railway Freight Traffic in Prosperity and Depression.* New York: National Bureau of Economic Research. Occasional Paper 5. 1942. 51 pp.

Institute of Actuaries. *The Institute of Actuaries Year Book 1941–1942.* Cambridge, England: The University Press. 1941. xli, 158 pp.

Johnston, Douglas T. & Co., Inc. *An Investment Policy for a Period of Economic and Social Change.* New York: Douglas T. Johnston & Co., Inc. 1942. 54 pp.

Kerby, C. Edith. *Manual on the Use of the Standard Classification of Causes of Blindness.* New York: American Foundation for the Blind, Inc., and the National Society for the Prevention of Blindness, Inc. 1940. v, 25 pp.

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. Some items of minor interest to statisticians have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications which are to be found in the *Revue de l'Institut International de Statistique, Journal of the Royal Statistical Society, American Economic Review, Population Index, Transactions of the Actuarial Society of America, The Record of the American Institute of Actuaries,* and *Sankhyā. The Indian Journal of Statistics.*—Editor.

Kuhlman, A. F., Editor. *The Development of University Centers in the South.* Nashville: The Peabody Press and the Vanderbilt University Press for the Joint University Libraries. 1942. 128 pp.

Kulp, C. A. *Casualty Insurance, an Analysis of Hazards, Policies, Companies and Rates* (Revised Edition). New York: The Ronald Press Company. xv, 680 pp. $5.00.

Lange, Oscar, McIntyre, Francis, and Yntema, Theodore O. *Studies in Mathematical Economics and Econometrics.* Chicago: The University of Chicago Press. 1942. 292 pp. $2.50.

Lindquist, E. F. *A First Course in Statistics.* Boston: Houghton Mifflin Company. 1942. xi, 242 pp. $2.50.

Lindquist, E. F. *Study Manual for A First Course in Statistics.* Boston: Houghton Mifflin Company. 1942. 117 pp. $1.00.

McGrane, Reginald C. *The Economic Development of the American Nation.* Boston: Ginn and Company. 1942. xii, 691 pp. $3.75.

McIntyre, Francis Edgar. *The Economic Pattern of Corporate Earnings: A Statistical Analysis.* Bloomington, Indiana: The School of Business, Indiana University. Indiana Business Studies, No. 25. February, 1942. 88 pp.

Marget, Arthur W. *The Theory of Prices.* New York: Prentice-Hall, Inc. Vol. II. 1942. xxv, 802 pp. $6.00.

Mitchell, William N. *Management's Adjustment to the Changing National Economy.* Chicago: The University of Chicago Press. Studies in Business Administration of the School of Business. 1942. ix, 61 pp. $1.00.

Negley, Glenn. *The Organization of Knowledge.* New York: Prentice-Hall, Inc. 1942. xvi, 373 pp. $4.00 (trade), $3.00 (school).

Nourse, Edwin G. *Between Automatic and Authoritarian Price-Making.* Washington: The Brookings Institution. 1942. vii, 38 pp. 25 cents.

Ohio State University. *Proceedings of the 1941 Ohio Conference of Statisticians on Business Research.* Columbus: Ohio State University Press. College of Commerce and Administration, The Bureau of Business Research. vii, 77 pp.

Pearson, Frank A., and Bennett, Kenneth R. *Statistical Methods Applied to Agricultural Economics.* New York: John Wiley & Sons, Inc. 1942. vii, 443 pp. $4.00.

Reed, Harold L. *Money, Currency and Banking.* New York: McGraw-Hill Book Company, Inc. 1942. viii, 522 pp. $3.75.

Seidemann, Henry P. *Curtailment of Non-Defense Expenditures.* Washington: The Brookings Institution. 1941. vii, 54 pp. 25 cents.

Seidman, Joel. *The Needle Trades.* New York: Farrar & Rinehart, Inc. 1942. xviii, 356 pp. $2.50.

Spiegel, Henry William. *The Economics of Total War.* New York: D. Appleton-Century Company, Inc. 1942. xiv, 410 pp. $3.00.

Stalson, J. Owen. *Marketing Life Insurance, Its History in America.* Cambridge, Massachusetts: Harvard University Press. 1942. xl, 911 pp. $6.00.

Warren, Robert, Jones, Homer, et al. *Financing the War.* Philadelphia: Tax Institute, University of Pennsylvania. 1941. ix, 357 pp. $2.50.

Woytinsky, W. S. *Three Aspects of Labor Dynamics.* Washington: Social Science Research Council, Committee on Social Security. 1942. xiv, 249 pp. $2.50.

Wright, David McCord. *The Creation of Purchasing Power.* Cambridge: Harvard University Press. 1942. xiv, 251 pp. $3.00.

Government Publications

United States Department of Agriculture. *The Competitive Position of Dairying in Southern New England,* by Herbert C. Fowler. Technical Bulletin No. 812. 51 pp. 10 cents.

Federal Housing Administration. *A Handbook on Urban Redevelopment for Cities in the United States.* 1941. ix, 105 pp. 15 cents.

Federal Reserve System, Board of Governors of the. *Federal Reserve*

Charts on Industrial Production. 1942. 233 pp. $1.00.

Federal Works Agency. *Second Annual Report Federal Works Agency 1941.* 1941. v, 473 pp. 65 cents.

The National Archives. *Second Annual Report of the Archivist of the United States as to the Franklin D. Roosevelt Library, Hyde Park, New York.* 1942. 19 pp.

National Resources Planning Board. *After the War—Full Employment.* January, 1942. 10 pp.

Securities and Exchange Commission, Trading and Exchange Division. *Report to the Commission by the Trading and Exchange Division on Secondary Distributions of Exchange Stocks.* 1942. viii, 78 pp.

Social Security Board, Federal Security Agency. *Old-Age and Survivors Insurance Statistics. Employment and Wages of Covered Workers 1939.* Bureau of Old-Age and Survivors Insurance. 1941. viii, 305 pp.

——— *Social Security Yearbook for the Calendar Year 1940.* Annual Supplement to the Social Security Bulletin. June 1941. xix, 347 pp. 70 cents.

——— *Employment and Wages of Covered Workers in State Unemployment Compensation Systems . . . 1939.* Bureau of Employment Security, Research and Statistics Division, Employment Security Memorandum No. 17. 1941. xiii, 404 pp.

United States Tariff Commission. *Cumulative Supplement to Changes in Import Duties Since the Passage of the Tariff Act of 1930.* 1942. 44 pp. 45 cents.

——— *Synthetic Organic Chemicals, United States Production and Sales 1940.* Report No. 148, Second Series. 1941. v, 64 pp. 10 cents.

——— *Wheat and Wheat Flour.* Report to the President. Report No. 145, Second Series. 1941. v, 38 pp. 10 cents.

State of New York. *Report of the State Commissioner of Housing to the Governor and the Legislature of the State of New York.* Legislative Document No. 20. 1942. 52 pp.

// JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 219 SEPTEMBER, 1942 *Volume 37*

APPLICATION OF STATISTICAL METHODS TO ORDNANCE ENGINEERING*

By Leslie E. Simon†

WHEREAS quality control may be regarded as beginning in 1924, it is principally the last decade that has seen the rise of the industrial statistician or quality control engineer. The men in this role remind one of a conversation between two men, one a close friend and the other a casual acquaintance of a third man of somewhat unusual attainments.

"You say Jones is graduate both of law school and medical college?" asked the acquaintance.

"Yes," said the friend.

"What is the result of such training?" continued the acquaintance.

"Well," said the friend, after a thoughtful pause, "all the lawyers call him Doc and all the doctors call him Judge."

For my own part, I deem it no discredit that engineers should regard me as a statistician and statisticians regard me as an engineer. I am neither. I feel that I undoubtedly am a professional soldier engaged for the most part in the specialized field of ordnance, and that the somewhat aloof position is not without its advantages. Thus the prides and prejudices often associated with one professional field do not cause me to hesitate to use the techniques of another if they are of practical assistance in attaining the objective of my problem: more and better arms for the soldier.

This anomalous professional situation appears to be general among quality control engineers, and the growing demand of industry for *engineers who have statistical training* leads to the belief that the hybrid engineer-statistician is going to become one of the greatest industrial and technological assets of this decade. However, formal training cannot fully prepare one for practical problems. The clearly defined procedures of the textbook and classroom may seem to be almost lost in

* An address given at the Graduate School, the U. S. Department of Agriculture, January 8, 1942, by authority of the Chief of Ordnance, and at the invitation of my friend Dr. W. Edwards Deming.

† Lt. Colonel, Ordnance Department, U. S. Army, Director of the Ballistic Research Laboratory, Aberdeen Proving Ground.

the pattern of everyday things. Frequently one must not only seek solutions through the use of methods which apply only approximately at best; but must, through alert observation of current work, recognize or create the opportunities for the use of statistical methods. Gold is where you find it, the old prospectors used to say. So it is with the applications of statistical methods. Opportunities for the application of at least some form of statistical methods undoubtedly abound in almost all fields, but they may be hidden under a mass of the drab details of practical work. I should like to illustrate the hybrid nature of applications of statistics and the absence of any formal method in their discovery with an example with which I have been closely associated. It would be pleasant if I could stimulate your interest by suggesting a thrilling intellectual field like that presented by Mathematical Statistics, but both candor and sound policy suggest that instead I give a warning of the unappealing simplicity and drab practicality of these applications by quoting a few lines from an author whose name I have forgotten.

> I walked along the ocean shore
> And came upon a jelly fish—
> A freshly landed, sadly stranded,
> Iridescent, smelly fish.

Simplicity of the control chart technique. As known and used up to the present, there is no statistical tool in the hands of the industrial statistician that is as universally useful and as practically powerful as the control chart. Its practical usefulness is due at least in part to its utter simplicity of application. Many a well planned scheme has failed because it was too complicated, but I have never heard of one's failing because it was too simple; and it is difficult to realize properly what a slight degree of mathematical methodology is promptly branded by industry as a mass of theoretical circumlocution. The control chart, when used with the average, $\overline{X}$, and the range, R, slays the hydra-headed dragon of complication; deflates the hobgoblin of fear-of-the-theoretical; and requires a minimum amount of labor in any large scale application of statistical technique.

The control chart's usefulness inheres not only in its simplicity of application but also in its simplicity of concept. Almost any engineer will tell you that when test results differ too much from expected results (expected results, in general, being based on experience), there is either something wrong with the product or the test. If the engineer is honest, he will likely reflect a moment and add, "However, it is very difficult to tell when results differ *too much.*" The control chart does just that—

it tells simply and clearly and in an economical way, just when one should take action on the grounds that a set of test results differs too much from its predecessors.

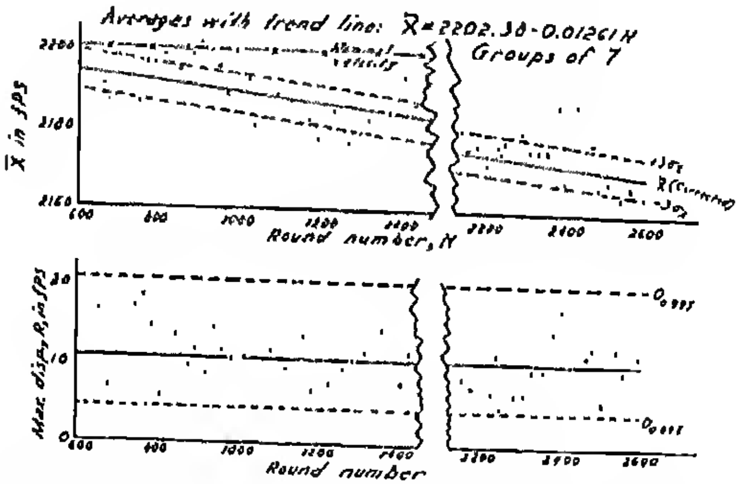
A new use for the control chart. The control chart has been applied for the most part to the control of manufacturing process and the judging or grading of the finished product with respect to quality. I believe that the Ordnance Department is the first to apply intentionally and systematically the control chart technique to the control of measurements. Ordinarily, one thinks of the arsenals as the manufacturers of defense products; but the Proving Ground turns out a product which is just as materially important as the munitions, and its product is *measurements.* If its flow of measurements is subject to assignable causes of variation, good munitions may be rejected and poor munitions accepted; if the range of variation of its measurements (accidental errors) is large, more rounds have to be fired in testing the product than if its measurements were more precise. In order to save valuable time and material, statistical methods are used in the control of the Proving Ground's product. Although I must omit all technical and military details, I believe I can still give you an interesting example.

In the testing of munitions, it is often necessary to fire the product in a gun in order to obtain an observable manifestation of the quality characteristic under consideration. Under these circumstances, the gun is not a gun in the usual sense, but a part of a system of measuring instruments. In scientific work, it is a cardinal principle that all measuring instruments must be calibrated. It is not so well recognized that such calibration should involve three considerations: accidental errors, semi-constant systematic errors, and mean constant errors. Together with the statistical method, I shall illustrate these considerations.

In the proof of ammunition, the necessity for calibration of the gun as a measuring instrument has long been recognized in an engineering sort of way. At the beginning of each experiment made with a test gun, a series of rounds (say 5 to 7) is fired with standard components. If the nominal muzzle velocity of the gun is 1800 fps (feet per second), and the average of the calibration series is 1750 fps, the gun is said to have an erosion correction of +50 fps, and 50 fps is added to all observed velocities in connection with that phase of the experiment performed on the component subject to test. That is, it is presumed that due to wear, the gun is firing low. Of course, the correction tomorrow may be +60, and the next day +40 (see Chart I). Occasionally, during the early life of a gun, a negative erosion correction may be encountered, but customarily that kind of correction used to be ignored, since it was presumed impossible for a gun to wear backwards.

When I came upon the scene, as a statistician I said to myself, "This is utter nonsense. These people are treating as a part of a systematic error what is really nothing but an accidental error of observation. The wear of a gun is not a discrete but a continuous function.[1] One should merely average one's calibration firings; keep a control chart (fitting the central line by least squares, if necessary); and as long as a single calibration series is within control limits, one should make a correction

CHART I
CONTROL CHART FOR VELOCITY OF A GUN
Averages and Dispersions for Groups of 7



which is equal to the difference between the nominal velocity and the central line of the control chart. This method is obvious and manifest."

However, as an engineer, I said, "Quiet! Say nothing about this yet, unless you want to damage your engineering reputation. Statistically, you may make a stupid statement because you are not sufficiently aware of the engineering considerations that underlie the phenomenon."

As an Ordnance Officer, I said in Al Smith's well-known phraseology, "Let's examine the record."

I got the records of firings reaching back into the years. I examined the lives of guns that had been fired to the end of their test lives. Chart I is a typical control chart resulting from such an examination. Whether I used the symmetric ±3 sigma limits[2] or percentage limits

[1] Neglecting, of course, discontinuities such as those that might result from periodic coppering and decoppering of the bore.

[2] For discussion of kinds of control chart limits see "Control Chart Method of Analyzing Data," Z1.2-1941, p. 11. The American Standards Association, 29 West 39 Street, New York. See also Leslie E.

(usually the 0.001 and 0.999 limits for $\overline{X}$ and the 0.005 and 0.995 limits for R), the results were always the same. The dispersions would show very satisfactory control. The averages invariably showed frightful lack of control, with from 10 to 40 per cent of the points outside of limits. Fitting the central line by least squares or allowing it to go down by steps through the life of the gun made little difference. Certainly external causes were being superimposed upon the random cause system which should have governed the behavior of the means. The contrast between the quite satisfactory control of the dispersions and the lack of control of the means made this obvious, but I had no improvement to offer on an existing system of doubtful validity until I could explain both why it was invalid and what should be done about it.

Naturally one suspects cyclical trends in data such as these. Statistical tests confirmed the suspicion of lack of normality. I tried to correlate the deviations of the sample means from the grand mean with the time of the year, the temperature, humidity, air density and various other phenomena without success. It is a widely accepted axiom that the statistical method merely detects trouble. I tried a large number of the tricks of the trade, but the statistical attack would not yield a solution of the causes of the trouble. Its identification and elimination were engineering problems. Here was freshly landed, sadly stranded, iridescent, smelly fish.

Engineering analysis of the record. One might well suspect that the standard components with which the gun was fired were not in a state of statistical control. However, ordnance products, especially those manufactured in the regular arsenals, tend toward a state of statistical control even when manufactured only under engineering control, perhaps because they are removed by several degrees of refinement from the initial raw material. This condition may be contrasted, for example, with the steel industry. Control also is fostered by the fact that the hazardous character of the material renders it necessary that it be scrupulously manufactured on a quality rather than a dollar basis. Furthermore, the most influential component in the calibration problem is the powder. I used to be closely associated with this component as Assistant Chief of Manufacture at Picatinny Arsenal, where it was made, and ran many quality control tests on it at that time. Hence, I believed the powder was controlled. However, conditions like those shown in Chart I make one doubt almost anything. I wondered if the broken sequences arising from the fact that only the calibration firings

Simon, *An Engineers' Manual of Statistical Methods*, John Wiley and Sons, New York, 1941, Chapters IV, VII and Appendix O.

are shown on the control chart, with the intervening service firings omitted, could have anything to do with it, so I decided to run control charts on some series of firings of the same set of components where no other components interrupted the series.

Series firings of the same set of components are almost always short, and are fired in a single day or two days. I gathered what I could. By diligent search of the records a considerable number of charts were made which were like Chart II except for length. Lack of control was almost never indicated either for standard components or even for non-standard components submitted for test.

Now there were enough pieces of the puzzle assembled to fit into an integrated whole. Firings show control within the day. Firings frequently show lack of control from day to day. Therefore, the lack of control rests not with the components and not with the gun but with the periodic introduction of a different systematic error in the process of making observations.[3] That is, there is (1) a mean error, or rather deviation from standard, associated with a short series which is due to wear of the gun (this causes the central line to slant), (2) a systematic error which is constant for a day or a set-up of apparatus, but which varies from day to day ($\overline{X}$ is temporarily shifted up or down, thereby causing points outside of the 3 sigma limits), and (3) there is the usual accidental or random sampling error (which may sometimes cause points to fall still farther outside the 3 sigma limits, or sometimes bring them in).

This analysis is in agreement with engineering considerations. The so-called accidental error is due to a combination of sampling fluctuations in the product (the standard components) and accidental errors of observation. The functioning of the systematic error is not so obvious, but it can easily be explained. Velocity is an indirect measurement which is inferred from the measured time-interval required for the projectile to pass from one screen erected along its trajectory to another. Having set up the screens and other necessary apparatus, these conditions remain relatively constant during an experiment. But, upon the repetition of the experiment (another standard powder firing) the screens are set a little closer or a little farther apart, the other necessary apparatus is slightly changed, and a different systematic error is introduced for that experiment. About 17 independent things affect this systematic error, sometimes called the error of the day.

Now one may talk freely about the trouble either as a statistician,

[3] For an earlier study see Student, "Errors of Routine Analysis," *Biometrika* XIX, pp. 151–164, 1927.

engineer, or professional worker in a special field; but the remedy for the trouble is not yet at hand.

The engineering and statistical analyses have now changed the original problem of "erosion" corrections and complicated it by logically injecting the issue of control. If observed average velocities are accepted without restriction, the manner of application of the "erosion" correction may be relatively unimportant, since a large error in the set-up of the measuring apparatus may result in an entirely invalid average velocity. Therefore, one must first seek criteria whereby he can judge whether the observed value should be accepted at all or whether action should be taken to discover assignable causes for variation in the measuring system, or perhaps even repeat the experiment.

An entirely satisfactory solution for this rather complicated problem is not likely, but surely statistical methods have something to offer which is better than unaided judgment. In the first place, we already know σ_a. If one knew the true average velocity for each of the respective set-ups, it would be a simple matter to calculate σ_s (the standard deviation of the semi-constant or systematic error). Then, by the rule of the sum of two variables, the total variation for observed averages of samples of n would be

$$\sigma_T = \sqrt{\sigma_s^2 + \frac{\sigma_a^2}{n}}.$$

However, the systematic deviations cannot be directly observed. Only n measurements of each respective deviate are known. Hence, were it not for the wear of the gun; i.e., the progressively shifting mean as shown by the trend line of Chart I, σ_T could be calculated directly from the observed means. Fortunately, there is a very simple method of eliminating the effect of a shifting mean, which has been used in exterior ballistics for a number of years. One merely calculates σ_T as

$$\sigma_T = \sqrt{\frac{\sum_{i=1}^{n-1} (X_{i+1} - X_i)^2}{2(n - 1)}},$$

instead of

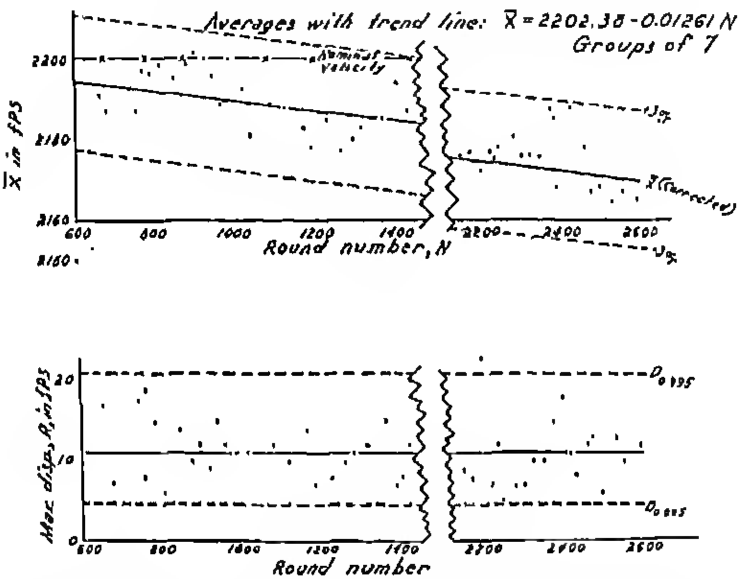
$$\sqrt{\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n}}.$$

Thus, one can compute σ_a, σ_s, and σ_T. One can now establish control chart limits for averages, from sample size n, at $\bar{X} \pm 3\sigma_T$. This control chart leaves some things to be desired, for from the viewpoint of systematic errors the sample size is one, and furthermore, from the viewpoint of systematic errors it falls short of being a control chart in the Shewhart sense, as it merely detects whether future measurements ap-

CHART II

CONTROL CHART FOR VELOCITY OF A GUN

Averages and Dispersions for Groups of 7



pear to be better or poorer than past measurements. However, it appears that this procedure certainly will not tend to indicate trouble when there is no trouble. This is healthful, for a system which cries, "Wolf! Wolf!", when there is no wolf will at once be regarded as ridiculous and discarded. It is essential, especially at the outset, that there be trouble, when it is so indicated, *and that it be found.*

When reputable engineers do work under carefully controlled conditions it is not rife with errors—neither is it entirely free of them. If the control chart procedure just outlined be applied to past data, certainly one cannot find the assignable causes of variation when such causes are indicated, since the measuring set-ups were dismantled long ago. However, one can observe whether the frequency with which trouble is indicated is reasonable. Let us apply the procedure to the data of

Chart I (real data so far as the statistical illustration is concerned, it has merely been altered so as to rob it of all military or engineering significance without prejudice to its illustrative value). Chart II shows this application. No point of those shown is beyond the $3\sigma_T$ control limits. (There was one out in the data not shown.) In like manner, the procedure was tried against very much more proving ground data before it was recommended for application.

I shall wish to say more about the use of this special procedure, but in the interest of coherence it must be observed that the original problem was the "erosion" correction. That problem has not been solved. The engineer said that the correction should be the nominal velocity minus the observed. The statistician said that it should be the nominal velocity minus the current value of $\overline{\overline{X}}$. At this point the issue became confused, for without some knowledge of statistical control, one knows little about $\overline{\overline{X}}$. The statistical problems are now fairly well cleared up. Are we now able to answer the question of the best value to take for the so-called "erosion" correction?

Yes, we can at least give an approximate answer, but first let us make several observations. If $\sigma_o/\sqrt{n}$ were very large as compared with σ_s (see Chart IIIa), certainly one should take the central line as the value of that day's velocity as suggested by the statistician. This is precisely what he had in mind when he suggested it. If σ_s were very large as compared with $\sigma_o/\sqrt{n}$ (see Chart IIIb), certainly one should take the point itself as the best estimate of the day's velocity as was the past practice. However, neither of these conditions obtains in practice, and in general the ratio of σ_s to $\sigma_o/\sqrt{n}$ lies between $\frac{1}{4}$ and 4 (see Chart IIIc).

If a point happens to fall within the $3\sigma_T$ limits but outside of the $3\sigma_s$ limits, we are almost certain that the observation has suffered from sampling fluctuations, for the $3\sigma_s$ limits mark the range of practically all true values of $\overline{X}$. Hence, we should certainly assume that the true value for the day is at least within the nearest boundary of the $3\sigma_s$ range. It may be well within the range; it may be directly on the central line. However, knowledge ceases with the boundary, and any advance within the boundary is an advance into the unknown.

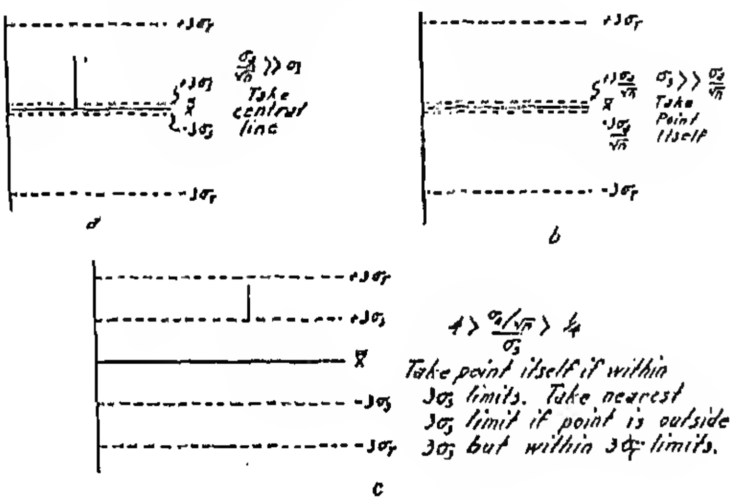
If a point happens to fall within the $3\sigma_s$ limits, I can see no reason for not accepting the observations as the true value of the velocity for that day.

There may be better rules than those shown in Chart IIIc. Certainly these are open to considerable criticism.[4]

[4] Since the date of this address, a solution has been obtained at the Ballistic Research Laboratory which minimizes the root mean square error of the estimated "true value for the day."

However, the important point of the whole standard powder procedure turns out to be the control chart technique; i.e., the rejecting of extreme and improbable values, and the identification and elimination of the causes that produced them. The $\pm 3\sigma_s$ band is in general relatively narrow, and the error committed is small even if the boundary of the band is assumed to be the true "velocity of the day," when actually the true value is the central line. This condition exists, because the really serious errors do not occur under the control chart technique. Hence, the injection of the issue of control was not really a digression

CHART III
CHOICE OF CORRECTION FOR "ERROR OF THE DAY"



from the subject of "erosion" correction. It is the major part of the solution. It is further obvious that increased accuracy can be obtained only through increased sample size, and that the velocity of the day may be known with whatever accuracy one chooses, at the expense of increased samples.

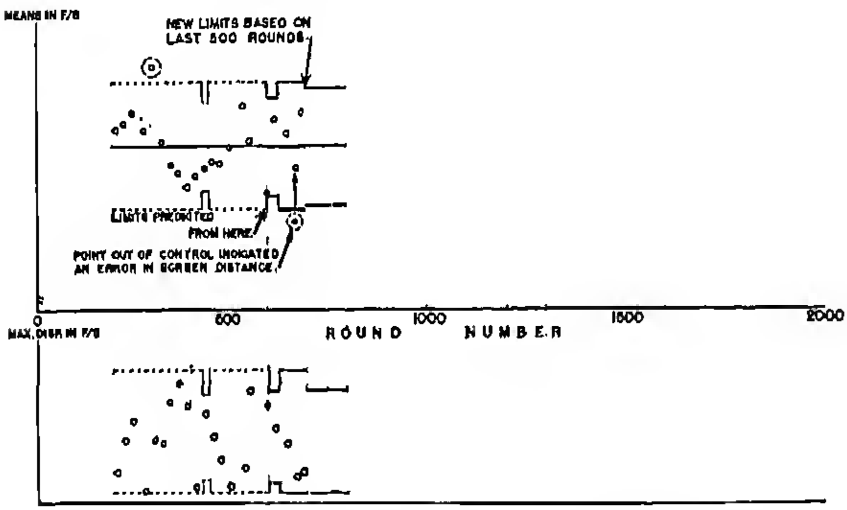
This may appear to be rather a simple problem with a simple answer, but things do not seem that way when one is working on them. What I have outlined represents part-time work for a year;[5] and whereas I feel that I should have made swifter progress, considerable experience in research and development work indicates that all new things—even new or only slightly different applications of old techniques—come slowly. If a man can make a substantial contribution to only one or two

[5] I am indebted to Dr. Walter A. Shewhart of the Bell Telephone Laboratories and Dr. W. Edwards Deming of the Bureau of the Census for helpful advice and consultation on this problem.

new things in a year, it is a considerable accomplishment. I make these remarks because one who is engaged in quality control work is apt to feel that his progress is slow, and subject himself to a considerable amount of harsh self-criticism, when actually he is making very creditable progress.

Of course one wonders how the procedure worked out in practice. Up to the present, it is working very well indeed, and the credit for its success rests for the greater part with Mr. David Kinsler of the Aberdeen Proving Ground, who had charge of its practical application. By virtue of good initial preparation, energy, resourcefulness, and persevering

CHART IV
POWDER CONTROL CHARTS



application, he ironed out some wrinkles in the plan, gained its acceptance on the part of operating personnel, and made several substantial improvements in its statistical and engineering features.

Charts like Chart II have now been kept on many guns, although the grand average is generally plotted horizontally by steps instead of on a slant. Chart IV shows the first working chart that was kept, and the first point which was out of control. It will be noted that a considerable number of points had to be accumulated, before control limits could be predicated. (At the present time approximate limits can generally be predicted at the outset.) During this period, a point occurred which was out of control and which, of course, was not recognized at that time. Work actually started at about round number 500,

and lack of control was indicated at round number 675. At that time we were running what I call "fini la guerre"; i.e., when a point goes out of control, everything stops until the trouble is found. The firing took place at noon, and it was 3:30 p.m. when we found the trouble. That first trouble just had to be found. It was an error in screen distance. That is, the two velocity screens were supposed to be 200 feet apart. An error of an even foot had been made in their measurement thereby resulting in a velocity determination which was ½ per cent too low. Since this measurement is ordinarily checked and double checked, trouble was not sought in that locality except as a last resort. However, it is hard to tell when and where people will make errors. The catching of this error surely put a feather in the cap of quality control. Care was taken to indulge in no recriminations or to place no undue blame on anyone for honest mistakes. Instead, everyone was encouraged to get behind the system and work for better measurements. At the last report I had on the system, 10 instances of lack of control had occurred and the assignable cause had been located in all instances except one. Of course, all instances of lack of control may not have been detected, but the errors could not have been large, for the control limits prohibit large errors. Hence, there is very strong assurance of high quality measurements from the Proving Ground.

I think also, there is an interesting corollary to the control program. During this emergency, proving ground activities have expanded enormously and many relatively untrained persons are charged with important measurements. A decline in standards is to be expected. Yet, under existing management and with the control chart technique, the charts show that the precision is no poorer than it was a year ago; in fact, it appears to be slightly better. This is a contribution of statistical methods to the war.

TESTS OF SIGNIFICANCE CONSIDERED AS EVIDENCE*

By Joseph Berkson, M.D.
Division of Biometry and Medical Statistics, Mayo Clinic

"After all, the higher statistics are only common sense reduced to numerical appreciation."—Karl Pearson.

There was a time when we did not talk about tests of significance; we simply did them. We tested whether certain quantities were significant in the light of their standard errors, without inquiring as to just what was involved in the procedure, or attempting to generalize it. In recent years tests of significance have been more broadly conceived as tests of hypotheses, and they have been generalized as t tests, F tests and certain amplifications of these, such as analysis of variance or of covariance. It is hardly an exaggeration to say that statistics, as it is taught at present in the dominant school, consists almost entirely of tests of significance, though not always presented as such, some comparatively simple and forthright, others elaborate and abstruse. Behind this is a doctrine of analysis that consists of setting up what is called a "null hypothesis" and testing it. Indeed, in this conception not only does this procedure characterize the method of statistics, but it is considered to be the very essence of all experimental science. In his well known book, *The Design of Experiments*, R. A. Fisher wrote, "Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis."[1]

What is this null hypothesis procedure? I quote from a recent text.[2]

> We have just set up the hypothesis that our sample of 900, which has a mean of 15,071 miles, is a random sample drawn from the population having a known mean of 15,200 miles. . . . Such a hypothesis is called a *null* hypothesis since our computations undertake to nullify it. The procedure may be summarized into three steps: (1) Set up the hypothesis that the true difference is zero. (2) Upon the basis of this hypothesis determine the probability that such a difference as the one observed might occur because of sampling variations. (3) Draw a conclusion concerning the hypothesis. If such observed difference could hardly have occurred by chance, we have cast much doubt upon the hypothesis. We therefore abandon the hypothesis and conclude that the observed difference is significant.

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 29, 1941.

[1] R. A. Fisher, *The Design of Experiments*. Ed. 2, London, Oliver and Boyd, Ltd., 1937, p. 19.

[2] F. E. Croxton and D. J. Cowden, *Applied General Statistics*. New York, Prentice-Hall, Inc., 1940, p. 310.

This I believe is a fair if abbreviated statement of the essential procedure as it is generally understood. If the experience at hand would occur only very infrequently in a given hypothesis, the hypothesis is considered disproved.

The argument has an apparent plausibility and for many years I adhered to it. However, set against experience with actual problems, reflection has led me to the conclusion that it is erroneous, and that a re-evaluation will lead to clearer comprehension in the application of tests of significance and also serve as a corrective of some of its misuses.

In the first place, the argument seems to be basically illogical. Consider it in symbolic form. It says "If A is true, B will happen sometimes; therefore if B has been found to happen, A can be considered disproved." There is no logical warrant for considering an event known to occur in a given hypothesis, even if infrequently, as disproving the hypothesis.

More to the present point, the argument does not seem to accord with what would be the mode of reasoning in ordinary rational discourse, nor with the rationale of usual procedures as they are observed in the scientific laboratory. Suppose I said, "Albinos are very rare in human populations, only one in fifty thousand. Therefore, if you have taken a random sample of 100 from a population and found in it an albino, the population is not human." This is a similar argument but if it were given, I believe the rational retort would be, "If the population is not human, what *is* it?" A question would be asked that demands an *affirmative* answer. In the null hypothesis schema we are trying only to nullify something: "The null hypothesis is never proved or established but is possibly disproved in the course of experimentation." But ordinarily evidence does not take this form. With the corpus delicti in front of you, you do not say, "Here is evidence against the hypothesis that no one is dead." You say, "Evidently someone has been murdered."

Nor do you find experimentalists typically engaged in disproving things. They are looking for appropriate evidence for affirmative conclusions. Even if the mediate purpose is the disestablishment of some current idea, the immediate objective of a working scientist is likely to be to gain affirmative evidence in favor of something that will refute the allegation which is under attack.

Does this mean that the application of tests of significance is in basic disaccord with rational scientific procedure? I am not sure. I think that there is a possibility of using them soundly, but the rule of inference on which they are supposed to rest has been misconceived, and this has led to certain fallacious uses.

Consider the objective of testing whether a distribution is normal. One could validly say, "If the distribution is normal and, the skewness of the sample, g_1, having been calculated, if a die of 100 faces, five of which are black, is thrown at random, a black face will occur only five times in 100." No one would suggest that the finding of a black face on a die following such a calculation is any reason for rejecting the null hypothesis that the distribution is normal. But when one says, "If the distribution is normal, a value of $g_1/S_{g_1} \geqq 1.96$ will occur only five times in 100," the finding of such a value for g_1/S_{g_1} *is* taken as reason for rejecting the null hypothesis. What is the essential difference between the two situations? Following the procedures which were outlined for dealing with a null hypothesis, one should reject the hypothesis that the distribution is normal on the finding of a black face, for it is surely an event rare in the circumstance of the distribution being normal. The difference appears to be that we recognize that if the distribution actually were *abnormal* (skew), the occurrence of a black face still would not be expected, but a large value of g_1/S_{g_1} *would be expected.* The latter constitutes evidence *in favor* of skewness. We may discern, as operating in the realm of tests of significance, a principle that I suggest is generally operative in scientific inquiry; it is this. The finding of an event which is *frequent* under a hypothesis H_1 can be taken as evidence *in favor* of H_1. If H_0 is a contradictory alternative to H_1 for which the event would not be frequent, then per corollary the finding of the event is, in so far, evidence in disfavor of H_0.

At this point I can imagine the question rising, "What difference does it make whether you say that you *reject* H_0 because for it the event is not frequent, or because you are *accepting* the alternative H_1 for which it is frequent?" To this the first answer must be that it would seem to be a sound idea to get one's head clear as to what are the principles on which one is really acting. If an event has occurred, the definitive question is not, "Is this an event which would be rare if H_0 is true?" but "Is there an alternative hypothesis under which the event would be relatively frequent?" If there is no plausible alternative at all, the rarity is quite irrelevant to a decision, and if there is such an alternative, the decisive question is, "Would the event be relatively frequent?" Secondly, the pursuit of a false principle for testing the null hypothesis will lead to false conclusions that will be avoided if one is consciously guided by the principle suggested here as being the correct one. I shall cite an example.

As an illustration of a test of linearity under the caption, "Test of straightness of regression line," R. A. Fisher utilizes data relating the

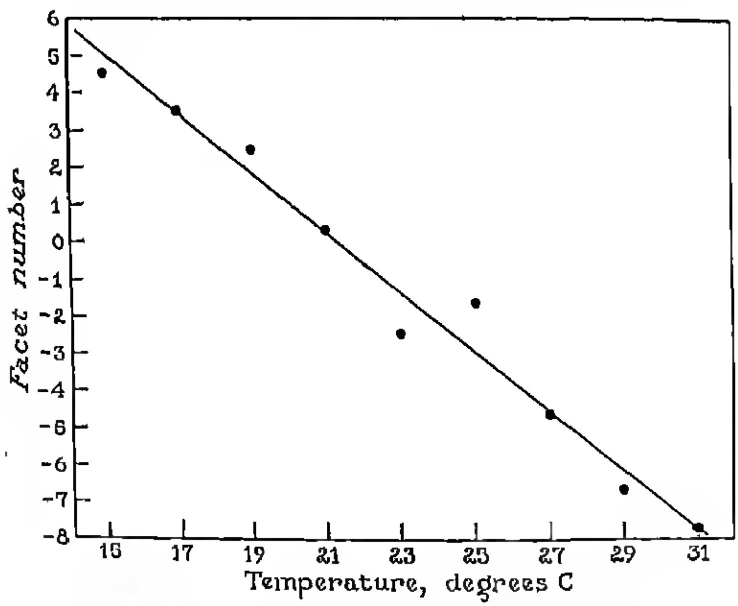
temperature to the number of eye facets of Drosophila melanogaster, the facet number being measured in factorial units. An analysis of variance procedure is utilized for the test and, the calculations having been made, Fisher says,

> The deviations from linear regression are evidently larger than would be expected, if the regression were really linear, from the variations within the arrays. For the value of z we have 1.2434 while the 1 per cent point is about .488. There can therefore be no question of the statistical significance of the deviations from the straight line . . . the departure from linearity was markedly significant.[1]

I have plotted the data of mean facet number in relation to temperature together with the least square line and they are shown in Chart I.

CHART I

MEAN NUMBER OF EYE FACETS OF DROSOPHILA MELANOGASTER RAISED AT DIFFERENT TEMPERATURES AND BEST FITTING STRAIGHT LINE BY METHOD OF LEAST SQUARES



Source: Data from R. A. Fisher, *Statistical Methods for Research Workers.* London, Oliver and Boyd, 1938, p. 260.

It was found by the significance test as applied that this regression was not straight, but on inspection it appears as straight a line as one can

[1] R. A. Fisher, *Statistical Methods for Research Workers.* Ed. 7, London, Oliver and Boyd, Ltd. 1938, pp. 259-265.

expect to find in biological material. What has betrayed the author is a faithful adherence to an unsound principle: to wit, reject the null hypothesis tested, in this case that the regression is linear, if the P of the test is small.

Let us consider the problem according to the principle advanced here. The event which has been found to have happened, in this case the small P, is to be considered as evidence in favor of any hypothesis under which it would be a frequent occurrence. Under what hypothesis would the P, considering its mode of calculation, be a frequent occurrence? If the regression were curvilinear, a small P is to be expected relatively frequently. In so far as this is so, a small P is evidence *in favor of curvilinearity* and because of this and *primarily because of this*, a small P can be considered evidence in disfavor of its alternative, linearity. But also a small P is to be expected relatively frequently if the regression is linear and the variability heteroscedastic; hence a small P is also evidence in favor of linearity plus heteroscedasticity. Or again a small P is to be expected frequently if the regression is linear and a value of the abscissal variate, in this case the temperature, is not constant but subject to fluctuation. And there may be other conditions which, with linearity, would produce a small P relatively frequently. The small P is favorable evidence for any or several of these. Which of these shall be taken to have been demonstrated by the evidence of the small P will have to be determined by other evidence, possibly other statistical tests. In this case my own judgment would be, not that the regression is nonlinear, but that the temperature has varied during each or some of the experiments. At least that would explain the small P.

According to what is advocated here, we cannot lay down any pat axiomatic rules such as "A very small P disproves the hypothesis tested," or "Equally, a very high P disproves the hypothesis," for it is not primarily the infrequency of the P which gives the finding its meaning. Each test will have to be examined and the circumstances in which it is applied will have to be examined, to find out, as best we can, whether any particular regions of P will occur relatively frequently in the case of an alternative to the tested hypothesis. There are situations in which a very large P will be frequent in an alternative, and in these circumstances, but *only in these circumstances*, a very high P can be said to disfavor the null hypothesis. I cite an example.

If with $(n+1)$ observations from a frequency distribution of a variate x the quantity $ns^2/\bar{x}$ is calculated, where $\bar{x}=\Sigma x/(n+1)$ and $s^2=\Sigma(x-\bar{x})^2/n$, it is known that the quantity is distributed in random samples as χ^2 for n degrees of freedom, if the distribution is Poisson.

Small values of P, say $P \leqq 0.05$, will occur with the small frequency of five times in 100. If, however, the distribution is what has been called supernormal, a distribution that is known to characterize certain physical situations, the variance σ^2 is greater than the mean μ, and in random samples large values of the quantity χ^2, and correspondingly low values of $P \leqq 0.05$ will be more frequent than five in 100. The finding of a $P \leqq 0.05$ therefore can be taken as preponderant favorable evidence for the super-Poisson, and hence as unfavorable to the null hypothesis tested that the distribution is Poisson. Similarly, if the distribution is Poisson, large values of P, say $P \geqq 0.05$, will occur with the small frequency, five times in 100. If, however, the distribution is Bernoullian-binomial or sub-Poisson, the variance σ^2 will be less than the mean μ, and small values of the quantity χ^2 and correspondingly large values of $P \geqq 0.95$ will be more frequent than five times in 100. The finding of a $P \geqq 0.95$ therefore can be taken as preponderant favorable evidence for the Bernoullian or sub-Poisson, and hence as unfavorable to the null hypothesis tested that the distribution is Poisson. Here then is a case in which either a very low value of P or a very high value can be considered as warrant for rejecting the null hypothesis. There are other such cases, but the rule is not general.

So much for the meaning of P's which are relatively frequent in the case of an alternative, and in so far, are evidence in disfavor of the null hypothesis tested. In the cases in which a very low P or very high P is evidence in favor of an alternative, what can we say of the finding of a middle value of P, say a P in the region 0.3 to 0.7? Statistical authors are not very clear about this. For the most part they merely confine themselves to statements that a low P disproves and one which is not low does not disprove. In some cases they say explicitly that a low P *disproves* but one which is not low does not *prove* the null hypothesis. What such a P should mean according to the principle advanced here is unequivocally clear. Since by definition such P's will occur frequently in the case in which the null hypothesis is true, the finding of one is to be taken as prima facie evidence *in favor* of the *null hypothesis*. That is in fact the way the statistician uses them, in contradistinction to the way he says they should be used when he describes the testing of the null hypothesis.

This was somewhat amusingly illustrated at one of our meetings. One of our most eminent members gave a paper presenting the application of the lambda test and used for illustration data designed to test a certain Mendelian hypothesis. The data having been examined and the test applied, a P of about 0.6 was found. "We can say therefore," he

remarked, "that the results substantiate the hypothesis." He applied the test illustratively to several other sets of data successively and getting a P of considerable size, each time he said, "The results therefore substantiate the hypothesis." When he was finished, an equally eminent mathematical colleague rose to object and said, "You cannot say that the results of the test support the hypothesis; all you are able to say is that they have not in these data disproved it." The most interesting part of the colloquy is that the first mathematician accepted the correction!

This I find is rather typical. In the abstract the mathematical statistician insists that a middle value of P only fails to refute the hypothesis, but if he is dealing with real data and gets interested in the physical problem in hand, he forgets his statistical principles and relapses to the rules of inference applied generally in such problems.

That statisticians with real problems in hand do interpret a middle P as positive support for the null hypothesis can be readily illustrated by innumerable examples to be found in the literature. I shall cite one that is in a field in which I once did some work. "Student," in his classic paper on the error of count with a hemocytometer,[4] used a series of data to examine whether the actual distribution in the hemocytometer followed the Poisson distribution, as it should on certain physical assumptions. He applied the Pearson chi-square test to a number of series and finding the P's taken together fairly large, he concluded that the distribution was sensibly Poisson, and that therefore the variability could be taken as the square root of the average count. If this positive conclusion in favor of the null hypothesis tested was not obtained from the relatively high P's, then his statistical work was entirely irrelevant. Other examples of the use by statisticians of relatively high P's for demonstration of the null hypothesis are easily found if one keeps a weather eye open for them.

When I say that a middle value of P is to be considered valid evidence in favor of the null hypothesis, I have by no means resolved all the pertinent questions that may be asked regarding it. I do not say anything has been "proved" or "disproved." I leave to others the use of these words, which I think are quite inadmissible as applying to anything that can be accomplished by statistics. All I say is that what we have is in the nature of positive supporting evidence. Whether the evidence is of sufficient weight to be convincing is another matter.

The development of what should be taken to affect the weight of the evidence is beyond anything I wish to undertake but a few pertinent

[4] Student, "On the Error of Counting with a Haemacytometer," *Biometrika* 5: 351–360, 1906–1907.

remarks I do wish to make. Whereas it can be said that the evidence provided by a small P correctly evaluated is broadly independent of the number in the sample from which it has been calculated, this is not true for such evidence as is provided by a P in the middle region, say 0.3 to 0.7. Consider Table I depicting the hypothetical results of a physician's judgments based on a serological test, designed to ascertain the sex of a fetus in utero. Examine experience 1, divest yourself of formal rules, and consider what would be your reaction. I think I can fairly guess that it would be something like this: "We cannot say anything from this experience: it certainly does not present any convincing

TABLE I
HYPOTHETICAL RESULTS: DETERMINATION OF SEX

Category	Experience 1			Experience 2		
	Total	Judgment of sex		Total	Judgment of sex	
		Correct	Incorrect		Correct	Incorrect
Expected by chance	10	5	5	1000	500	500
Physician's judgment	*10*	*6*	*4*	*1000*	*505*	*495*
P	0.38			0.38		

evidence that the physician can discriminate between the sexes. But I should not want to say either that he cannot discriminate. The experience is too small for any conclusion." With experience 2 I think you would say, or at any rate I should: "There is no question in my mind; quite evidently the physician does not possess any ability to discriminate by this serological test between the sexes. The experiment is quite large enough, and if he could discriminate to any significant degree we should see it in the results, which we do not."

Now for both experiences, the P, which is the probability of obtaining by chance as good a result as the one obtained, on the null hypothesis that the probability of either sex is a half, is the *same*, *namely*, 0.38. But the experience 2, being based on large numbers, is convincing positive evidence of the truth of the null hypothesis within practical limits. I do not intend to attempt to analyze what is the justification for the added conviction provided when the numbers are large, beyond suggesting that it has the same basis as what has been argued here is the general principle of inference which is operative throughout. When the numbers are small, a middle P will occur with considerable frequency if the null hypothesis is true or if an alternative is; with large numbers such a P will occur frequently in the case of the null hypothesis but not

in the case of a practical alternative. Hence with large numbers, a middle P provides probative evidence in favor of the null hypothesis.

Here we have disclosed one fundamental weakness in the position of those who contend that small samples can be effectively utilized in statistical investigations if the calculations of the P's are correctly made. If it were a fact that conclusions are drawn only when the P is very small and the null hypothesis disproved, then so far as concerns the main considerations here developed, there would be a certain validity to this view, for small P's are more or less independent, in the weight of the evidence they afford, of the numbers in the sample. But if actually it is the fact that conclusions will be drawn from P's which are not small, then only very considerable numbers in the sample are reliable.

If a test for the difference between means has yielded a large or middle P, it does not merely fail to disprove the null hypothesis that the true means are equal; it furnishes *affirmative evidence* that the means *are* substantially equal. If the numbers on which the test is based are large, the evidence will have convincing weight; otherwise not. Contrariwise a low P points affirmatively toward the alternative that the means are unequal. It is the merit of some kinds of tests that they indicate unequivocally the specific alternative toward which they point.[5] Such are tests for the difference between means or the difference between variances or tests for skewness. Other tests such as the frequency χ^2 or some applications of the analysis of variance do not have this characteristic. In Table II is presented an experience of mortalities following certain operations with and without the use of a vaccine for the prevention of peritonitis. Four tests are given for the "null hypothesis" that the true mortality rates are identical for patients with and without vaccine: (1) the probability of getting as many differences in the favorable direction as found; (2) the appropriate P for the χ^2 test of the four-fold table constituted by the totals; (3) the Fisher test of combining the value of $\chi^2 = -2\, ln\, P_{\chi^2}$; (4) the summation of the χ^2 and degrees of freedom for the separate operations. The resulting P's are considerably different. In terms of the usual rationalization, each of these tests is equally valid for testing the null hypothesis. If the null hypothesis were true, that is, if the vaccine were ineffective and the mortality for any operation were the same whether the vaccine were used or not, the appropriate limiting value of each test function would

[5] Elsewhere I have suggested that these tests are ones which in principle can be stated alternatively and equivalently in terms of an estimate and its confidence limits. Joseph Berkson, "Comments on Dr. Madow's 'Note on Tests of Departure from Normality' with Some Remarks Concerning Tests of Significance." This Journal 40: 539–541, December 1941.

occur only infrequently—one just as infrequently as the other. But the tests are differently sensitive to the presence of different *alternatives*. In terms of the Neyman-Pearson formulation they have different powers for any particular alternative, and hence are likely to give different results in any particular case. How blind is the procedure of doing some test of significance, when there is no knowledge at hand as to whether it is likely to show a significant result or not show one, no matter how importantly different the facts may be from the hypothesis tested. The importance of this consideration is underscored when we

TABLE 11

MORTALITY RATES FOR OPERATIONS WITH AND WITHOUT USE OF VACCINE: TESTS OF SIGNIFICANCE OF DIFFERENCES

Type of operation	Vaccine			No vaccine			Mortality difference, per cent
	Operation	Hospital deaths		Operations	Hospital deaths		
		Number	Per cent		Number	Per cent	
A	107	2	1.9	142	4	2.8	−0.9
B	28	3	10.7	60	9	15.0	−4.3
C	21	3	14.3	34	5	14.7	−0.4
D	21	4	19.0	34	8	23.5	−4.5
E	47	3	6.4	45	4	8.9	−2.5
F	21	1	4.8	26	2	7.7	−2.9
Total	245	16	6.53	341	32	9.38	−2.85

Test	P
1. Signs	0.016
2. Total difference mortality	0.11
3. Combination of P's—Fisher	0.01
4. Summation of χ^2 and D.F.	0.08

realize that in practical applications the failure to show a significant result will be taken to corroborate the null hypothesis. It is an important but neglected task of mathematical statistics to investigate what alternatives are particularly pointed to by specified findings with different tests.

I should like to see the development of investigation of the finding of middle P's. I am not ready to say what this should be or just what it would lead to. But this is an example of what I mean. With the development that we now have, which emphasizes the low P's, we find such statements as the following in the literature, and it is typical of the essential procedure in many fields in which statistical tests are applied. A standard curve for estimating dosage from mortality has been established with its confidence zones, from a first set of data. A set of data for another drug is to be used for estimating the potency of a second drug.

But realizing the possibility that the standard curve may not be applicable any more, the author counsels the use of some controls to see whether the standard curve still applies for the first drug. He says, "When the controls have been shown to agree with the standard of the regression line by the appropriate χ^2 or t test, the first curve can be used." Now what is meant by this is that if the test does not show a low P, the curve can be used, which is to say that if the test shows a middle P, the curve *will* be used. It should be clear on consideration that if there is a real discrepancy of a given size between the present conditions and the curve, a P which is not low will result with small numbers, while with the same discrepancy a low P will result if the numbers are large. The use of the suggested rule could easily be disastrous if drugs were standardized on the basis of it and small numbers were used. Investigation should be made which could result in a rule not such as just given, but rather of the following kind: "If the control is tested with data including so and so degrees of freedom and if the test results in a P of this amount or higher, the curve may be accepted as stable."

THE STATISTICAL WORK OF THE LEAGUE OF NATIONS IN ECONOMIC, FINANCIAL AND RELATED FIELDS

By Charles K. Nichols
Economic, Financial and Transit Department, League of Nations

It will come as a surprise to some to learn that the Economic, Financial and Transit Department of the League of Nations is maintaining the bulk of its work in Princeton, New Jersey, with certain basic tasks still being executed in Geneva. It is a tribute to the perseverance of its members as well as to the foresight and generosity of the member governments still contributing to the support of the League, and to certain private institutions[1] in the United States. Largely, the activities are a continuation of the vital fact-finding functions performed in Geneva, but in addition special studies on issues likely to be controverted in the post-war period are being prepared.

The chief consideration governing the decision to maintain this organization seemed to be that its efforts would be most useful in settling the great problems following the peace. The world then will be vastly different, and without up-to-date information the difficulties of making sound decisions in economic matters would be increased. Furthermore, what was learned during the inter-war period can be evaluated and crystallized into a positive plan for economic collaboration in the future. Also, the net-work of contacts and associations with government personnel can be at least partially maintained so that in the post-war period a more effective job of collecting data and promoting understanding in an informal way can result.

Before discussing the history of the economic and financial work it might be wise to recall briefly the wide scope of problems dealt with by the technical sections of the Secretariat assisted by advisory expert committees. Health, child welfare, opium control, housing, nutrition, land and sea communications besides economic and financial questions were carefully investigated in order to furnish the Council and the Assembly of the League, as well as the governments of the several states, accurate and meaningful information. In these tasks some of the Sections created a respectable body of statistical data, and *inter alia*

[1] The Department came to Princeton in September 1940 in response to a joint invitation extended to the Technical Services of the League by Princeton University, The Institute for Advanced Study and the Rockefeller Institute for Medical Research. The facilities of the two first mentioned academic institutions are being utilized for the work now carried on in Princeton and the Rockefeller Foundation is helping through a liberal grant.

exercised considerable influence on the methods employed for their collection and publication in the separate countries.

HISTORY OF THE ECONOMIC AND FINANCIAL WORK

Almost at the outset the Council of the League saw the necessity of maintaining a Committee for the consideration of economic and financial problems. Accordingly, as an outcome of the Brussels Conference in 1920, representatives in close touch with economic and financial ministries or central banks in many countries were chosen to meet regularly and take up questions referred to it by the Council. The early work centered upon reconstruction, largely in financial matters, and disorganized and newly-formed states were afforded invaluable assistance in reorganizing and organizing their financial and monetary affairs. Throughout the twenties this work was a dominant interest. Later, in 1935, a Fiscal Committee was formed to handle special problems of taxation and public finance. Gradually the Economic and Financial Committees became more active with inquiries into matters of industrial, economic and commercial importance and it was undertaken in diverse ways to bring governments together to discuss commercial policy. The World Economic Conferences of 1927 and 1933 were notable as efforts of these Committees to promote the ratification of agreements to improve international economic relations. These efforts were continued throughout the inter-war period but the emphasis shifted somewhat from attempts to secure ratification of general agreements to the procedure of promoting more limited accords among nations of similar interests, and also toward endeavors to influence national policies through focusing attention on problems common to many regions. Thus, by designating a sub-committee on the Mitigation of Economic Depressions, world experience was pooled and expert appraisal was given to the relative effectiveness of different measures introduced.

This work on depressions is continuing, although most of the Committees' activities have been suspended. Their past labors are available, however, in reports and memoranda[2] which afford to economists and statesmen a body of factual information and a historical record to serve as a guide to the formulation of international economic policy in the future.

The Economic and Financial Committees were appointed by the Council of the League, and were composed of civil servants and bank-

[2] See for instance the two Tinbergen reports on *Statistical Testing of Business-Cycle Theories* (L. of N. publications 1938.II.A.23 and 1939.II.A.16) and Professor Haberler's *Prosperity and Depressions* (L. of N. publication 1939.II.A.1).

ers. Their reports were prepared for submittal to the Council and thence were transmitted to governments. Distinct from these Committees has been the Economic and Financial Section of the permanent Secretariat, which embodied the important Economic Intelligence Service to be dealt with below. The Economic and Financial Section, made up of salaried experts engaged by the Secretary-General, was formed at the very first and has functioned throughout with a consistent purpose but a flexible program. Largely anonymously, its members have been responsible for the collection and analysis of economic information for the use of the Economic and Financial Committees, business men and bankers, economists, journalists and statesmen. Its purpose has been to observe and interpret world economic and financial events, and record a chronology of them in serial publications. In this task one of the chief objects has naturally been to assemble statistical information, which brings us to a consideration of the vast work which has been quietly carried on by the body now operating in Princeton, and currently designated as the Economic, Financial and Transit Department.

THE STATISTICAL WORK

As early as 1920 the League Council took up the question of economic, financial and social statistics through the instrument of an International Statistical Commission composed of members of such existing organizations as the International Institute of Statistics, the International Labor Office, The International Institute of Agriculture, and the International Chamber of Commerce. The Commission recommended that existing organs separate from the League should be utilized to furnish statistical information. A minority report, however, suggested also that a permanent organ be set up in the Secretariat for this task, and such became the case. It became then the duty of the Economic and Financial Section to institute a service to collect and publish economic and financial statistics. All fields were to be covered except agricultural and social statistics which the International Institute of Agriculture and the International Labor Office could furnish. It was also deemed advisable to consult the International Institute of Statistics thereby drawing on the accumulated technical knowledge of that body's experience in statistical theory.

Confronted with the task of collecting such diverse information from political divisions using widely different procedures, it was soon seen that international comparisons could only be made if the national statistics were available in comparable form, and represented not dissimilar phenomena measured by comparable methods. At first the work

was confined to the publication of the *Monthly Bulletin of Statistics*, and to the preparation of memoranda on special problems of trade, tariffs and finance. This brought the Section in close contact with government statistical offices and offered an opportunity to suggest methods for improving national statistics and for presenting them so as to obtain better international comparability. It was necessary, however, constantly to suggest and request, and without authority to insist, the work could not progress as desired. Nevertheless, it was possible in 1927 to commence publication of the *Statistical Year Book* embodying data included in the *Monthly Bulletin*, but enlarging on them.

It appeared, after several years' effort to induce governments to cooperate, that some official means would be necessary to secure this cooperation. Hence, in 1927 the Economic Committee recommended to the Council that a conference of government representatives be called to conclude an International Convention on Economic Statistics. Accordingly, in 1928 delegates from 40 countries and representatives of several Institutes concerned with international statistics were assembled to consider the agenda prepared by the Economic and Financial Section. The resulting Convention was ratified by 26 states and bound the Contracting Members to publish certain classes of economic statistics according to principles agreed on at the Conference and incorporated into the Convention.

This represented an unique step toward international cooperation in technical fields and furnished some official backing for the task of the Economic and Financial Section. And in fact it became less difficult from the date of that Conference to obtain statistics compiled according to similar principles and published in comparable form.

THE COMMITTEE OF STATISTICAL EXPERTS

Perhaps the most important result of the Conference was the creation of the Committee of Statistical Experts, provided for in Article 8 of the Convention. To this body, appointed by the League Council and constituted of specialists chosen on the basis of special competence, were delegated the tasks of further developing sound principles for the compilation of national statistics, and of examining the problem of international comparability.

Working in close connection with the Economic Intelligence Service, the Committee of Statistical Experts met in 1931 and yearly from 1933 to 1939 and succeeded in considering virtually the entire field of economic statistics, and from 1935 managed also to inquire into some problems relating to financial data. The results of their work are presented

in full reports on each session. Partial reprints containing the Committee's main recommendations on specific topics are available in a series of *Studies and Reports on Statistical Methods.*

The work of the Committee is so extensive it is impossible to cover each phase of it here. However, some indication of its importance can be realized by considering, as an example, what was done in the field of statistics on international trade. The Committee established by the middle of the thirties what is known as the *Minimum List of Commodities for International Trade Statistics.* Many countries, adhering to the convention of 1928, agreed immediately to use the list, while others followed in a few years. By 1939 it was possible for the Economic Intelligence Service to publish trade statistics for about 30 countries which had been compiled according to the minimum list. Thus, for the first time, truly comparable trade statistics were at the disposal of statesmen, economists and business men.

Likewise, with indices of foreign trade and industrial production, statistics on the gainfully-occupied population, housing, capital formation, balances of payments, tourism and several other topics the Committee has established principles which governments have shown increasing willingness to apply by adapting their national statistics to the basic recommendations of the Committee or by publishing supplementary tables drawn up in conformity with its Standard Classifications for purposes of international comparison.

The work of the International Labor Office deserves mention here as its activity in the field of social statistics complements that of the Economic Intelligence Service, and the two bodies have been closely associated. By similar methods the Labor Office has promoted the wide use of sound principles in compiling social statistics and has influenced the publication of internationally comparable information. The Economic Intelligence Service relies upon data collected by the International Labor Office in the fields of employment, unemployment, wages, migration, etc.

PUBLICATIONS

In its current statistical work the Economic Intelligence Service has naturally benefited from the advice of the Committee of Statistical Experts. The results are evident in the improvement and extension of the publications of the Service. The *Statistical Year Book* and the *Monthly Bulletin of Statistics* have been described as source documents designed for government and public use. Their chief advantage is the convenience of finding in a single volume information for every area of the world; and because of this convenience governments are spared considerable expense.

The two publications carry similar tables although the *Year Book* is more extensive. It includes data on population by age groups and occupation; births, marriages, and deaths; mortality rates, expectation of life, fertility and reproduction; employment and unemployment, and wages. The purely economic information deals with statistics of production, indices of primary and industrial output, and stocks; data on transport; and data on trade. Financial statistics cover currency, bank deposits, budget accounts, public debt, prices, exchange rates, discount rates, bond yields and capital issues.

Other volumes are based on the immense amount of statistical information collected by the Intelligence Service, but are also interpretive. *International Trade in Certain Raw Materials and Foodstuffs, World Production and Prices, Review of World Trade, Balances of Payments, Money and Banking* are all included in this group. The *World Economic Survey,* published each year since 1932, is devoted entirely to summarizing important economic happenings throughout the world. The first issue grew out of a special study on economic depressions conducted in 1931 and 1932, and the yearly issues supply a coherent record of economic events throughout the thirties.

These special memoranda are important in particular because they are the first attempt to use the raw material of national statistics to make a new kind of raw material for the study of world problems. In each sphere the volumes have drawn together data and compiled indices for the summarization of the economic trends of the *world as a whole.* These documents are accepted as important contributions to the study of world economic problems, and represent an unique methodology.

It is also significant that the entire range of economic phenomena is scrutinized by a unified group of economists. World problems are thus brought into full perspective by a comprehensive and coherent body of data of which the parts are interpreted by competent experts.

THE WAR AND AFTERWARDS

In a wartime world the activities of the League naturally have been reduced. The importance of maintaining the technical work has been recognized, however, and the Economic Intelligence Service continues to function. The *Year Book* and the *Monthly Bulletin* are being published in spite of the increased difficulty of obtaining information. The regular works on trade, production and prices, banking, finance, etc., have been replaced by briefer chapters on these topics in the *World Economic Survey.* The last *Survey,* issued in 1941, deals of necessity largely

with problems of war economics, as will the issue to be published in the fall of 1942.

The work of the Department relative to problems covered by documents such as the *Review of World Trade* and *World Production and Prices* has not ceased with the cessation of these publications, however. Data are being assembled and special memoranda such as *Raw Materials and Foodstuffs*, published in 1940; *Europe's Trade*, 1941; the *Network of World Trade; War-Time Rationing and Consumption;* and *Money and Banking 1940/42* (the last three recently published or to be published shortly) all cover significant happenings in the respective economic and financial spheres. They contain statistical materials and represent authoritative international sources of economic and financial data carefully sifted and digested.

Most of these documents form part of a new set of studies concerned with post-war reconstruction. By analysis of the problems and economic events leading up to and following the first World War, it is being attempted to anticipate, also in the light of information collated by the Economic Intelligence Service concerning trends during the inter-war period, what measures will likely be most effective in preventing a recurrence of the difficulties following the last war. In this work the statistics assembled by the Intelligence Service are performing the function they have been assigned—namely, they are the base upon which the analytical work is grounded. In this respect, then, the labors of this League organization to improve national statistics and to secure international comparability may prove valuable in settling the reconstruction problems which will be faced by a post-war world.

It is probably true that an orderly international system does not rest entirely upon such technical achievements. The fulfillment of psychological and deeper social needs is no doubt more fundamental in the process of attaining peaceful international relations. However, in critical periods the existence of a workable technique may be of first importance in promoting the greater end. At the next world peace conference when the idea of an international organization will undoubtedly enjoy unprecedented popularity the presence of a coordinated plan for economic cooperation may prove decisive.

THE MARKET FORECASTING SIGNIFICANCE OF MARKET MOVEMENTS

By L. C. Wilcoxen

As the seer studies his crystal, so too thousands of traders study their charts of the price movements of both stocks and commodities. They are the courageous exponents of the philosophy that, "If you want to know what the market is going to do, study the market." It is the purpose of this study to analyze the actions of certain markets in order to determine their inherent forecasting possibilities.

To establish the characteristics of market behavior, it is necessary to select moving definite time intervals of market action and to determine the probable actions in the immediate subsequent moving intervals. This statistical analysis yields specific "forecasting" criteria for these specific time intervals, which may be tested comparatively. From these results the criterion for maximum trading profit may be determined and its possibilities appraised.

The raw data employed in this study are the daily "highs" of the U. S. Steel Common stock; wheat May futures and cotton October futures. The "highs" were selected arbitrarily rather than the "lows," which would have served equally well. Daily figures were chosen to permit analysis, based on short time intervals, to give the greatest volume of data and to make possible more precise testing of the criteria than would be possible, for instance, by the weekly "highs."

The question of what subsequent market action will follow prior market action may be expected to depend upon the lengths of the prior and subsequent intervals. The amplitude of the subsequent market movement may be expected to bear some relation to the amplitude of the prior movement. Thus, this study begins with the selection of prior intervals, subsequent intervals and the statistical determination of the relative market movements.

The first of the pertinent intervals for which U. S. Steel stock was analyzed is the 1-1, that is a one-day-prior interval with its subsequent one-day interval. Other intervals are designated as the 2-2, 4-4, 8-8, 16-16, 32-32, 48-32 and 64-32, each pair of figures representing the prior and subsequent intervals in market days.

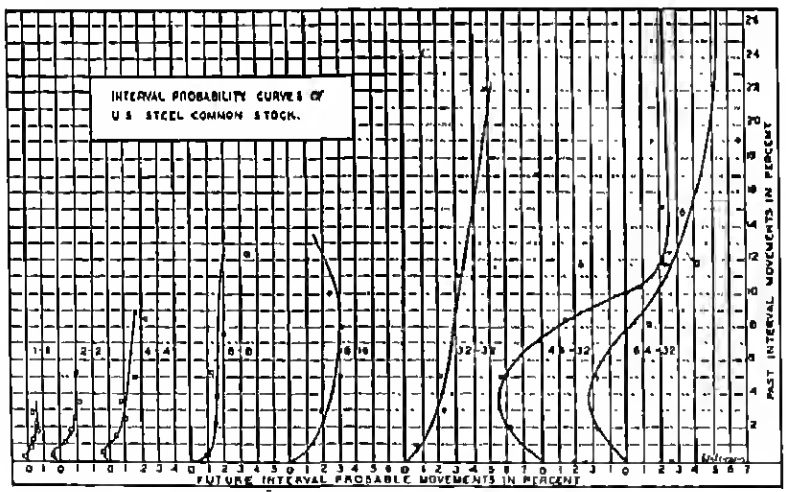
The data selected for the purpose of this study were the daily "highs" of the U. S. Steel Common stock prices from January 1, 1922, to January 1, 1932. This stock was deemed suitable because of its broad and continuous market action throughout this period, during which time it

was reasonably free from so called "manipulation." In this ten year period there were approximately 2900 market days.

First the percentage change for each of 2900 market one-day intervals (day to day) was determined. From these a series of frequency charts was made for successive brackets of percentage change from the prior day (one-day-prior interval). For instance, a frequency chart was made of all one-day changes following one-day percentage changes of from zero to 0.5 per cent. Percentage movements which were in the same direction were considered positive and those in the opposite direction, negative. From the frequency charts the probable error was computed from the formula $.6745\sqrt{\overline{x^2}}$. In this study the resulting probable error of each case is considered as the forecasting criterion.

For these one-day intervals, seven brackets of prior interval percentage changes were selected and frequency curves were made for each. The probable errors are spotted on Chart I.

CHART I



A figure has been computed for the probability of the total observations and is noted in Table I. The value of this figure, which will be designated as the Summary Probability Characteristic, is .286 per cent. It will appear later that this characteristic is important in making comparisons with the forecasting values of different criteria.

In Chart I, a probability curve is shown fitted by eye to the computed points. It is significant to note that one-day market movements of less than 0.7 per cent indicate probable reactions the following day.

Movements of greater amounts indicate that on the following day the movements will be in the same direction. In the aggregate the more numerous minor reactions practically counterbalance the movements in the same direction. In spite of this fact, the summary probability characteristic, which is based on the dynamics of the movement (X squares instead of X's) is of appreciable magnitude and in the same direction.

In the same manner the data for various intervals have been analyzed. Curves are fitted to the computed points for intervals of 2-2, 4-4, 8-8, 16-16, 32-32, 48-32 and 64-32 and are included in Chart I.

Note the change in the characteristics of the successive curves. The 1-1, 2-2 and 4-4 are similar in that minor movements during the prior

TABLE I

THE SUMMARY PROBABILITY CHARACTERISTICS FOR U. S. STEEL COMMON STOCK, WHEAT MAY FUTURES AND COTTON OCTOBER FUTURES FOR VARIOUS INTERVAL PROBABILITIES, USING THE FORMULA $.6745\sqrt{x}$

Interval	For U. S. Steel Common Stock S.P.C.	For wheat May futures S.P.C.	For cotton October futures S.P.C.
1–1	.280	.431	
2–2	.455	.554	
4–4	.783	.004	
8–8	1.414	.074	1.14
16–16	2.24	2.00	2.43
32–32	3.32	−2.04	3.76
48–32	2.08		
64–32	2.34		

interval are followed in each case by probable reversed movements. For each of the next several intervals, there appears a change in which the market action of all amplitudes is followed by market movements continued in the same direction. This is true for the curves 8-8, 16-16 and 32-32. Then on the next two curves—the 48-32 and 64-32—intervals revert back to the type of the three minimum intervals, that is, the 1-1, 2-2 and 4-4 curves.

This is truly interesting. It shows that for the 1-1, 2-2 and 4-4 periods, a minor cycle exists with amplitudes of something less than one per cent. Thus, while true, the fact is of trifling significance.

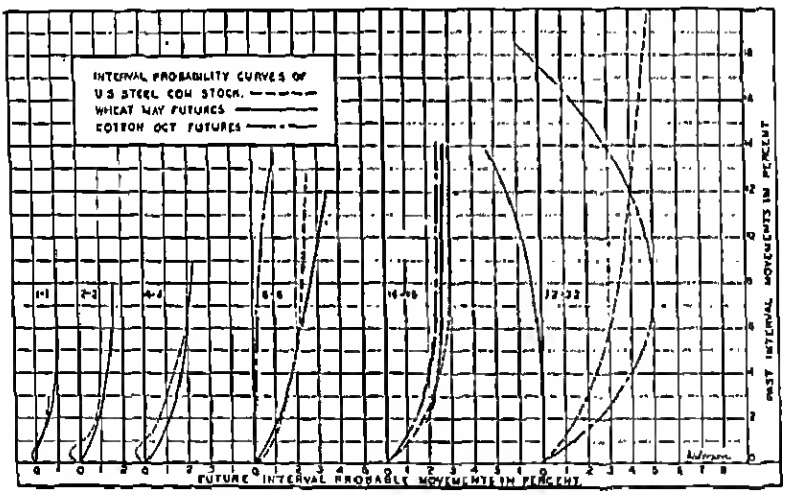
The reactive cycle does not again begin to appear until the intervals of 48-32 and 64-32 are reached when it is again qualified. In these cases only market movements of less than 8 and 12 per cent forecast a reversal (i.e. indicate a probable negative future movement), while those of greater amplitudes forecast a continued movement in the same

direction. Here, as in the 1-1, 2-2 and 4-4, the reverse movements practically counterbalance the continued movements. How important this defined movement is will be brought out by further analysis.

The principle by which the forecasting characteristics of U. S. Steel Common stock has been analyzed, undoubtedly has wide application. It is used here to further the understanding of free markets.

The characteristics of the wheat May futures market were investigated, using for data the daily "highs" for the nine years January 1, 1921, to January 1, 1930, inclusive, in which there were approximately 1900 market days. The forecasting curves, for several intervals analyzed are shown in Chart II, where they are compared with the curves of the U. S. Steel Common stock.

CHART II



In the same manner the cotton October futures market was analyzed, the data used being the daily "highs" for the ten year period of January, 1921, to December, 1930. In this period there were approximately 1850 market days. The curves for the intervals 8-8, 16-16 and 32-32 are also shown on Chart II. The important point to notice is that for short periods the forecasting curves of the several markets are markedly similar. It is quite possible that this is a general characteristic of all free markets.

Attention has previously been called to the Summary Probability (Forecasting) Characteristic of U. S. Steel Common stock for the 1-1 intervals. This characteristic, which it should be recalled, is the

weighted average of points on the probability curve, has been computed for steel, wheat and cotton for various intervals in each. For comparison, their amounts are shown in Table I.

Two things are strikingly brought out. First, the fact that, for the shorter intervals, the values of the characteristics are approximately the same for the different markets. This may indicate that fundamentally all markets are governed by the same law. Second, that for U. S. Steel Common, the value of the characteristic decreases after the 32-day-preceding interval, an indication that this may be the limiting period for which such criteria should be used for practical forecasting.

The various probability curves are statistically correct for anticipating the probable stock market action of U. S. Steel Common stock, and as they are specific, it is possible to test their comparative merits. Given the market action up to any date, current or past, it is possible to consider the probabilities as indicated from each of the curves as exact forecasts. The day to day progressive forecasts through any past period will give the same results as if the forecasts were made currently. Obviously the stock should not be bought or sold unless the probable movement is greater than the round turn cost of completing the transaction. It is equally obvious that once a commitment has been made, it should be maintained until the forecast is reversed again to the extent of the round turn cost.

The present (1042) percentage round turn cost is variable, decreasing from 4.4 per cent, when the price is $10 a share, to approximately 0.4 per cent when the price is $300 a share. These are the extreme percentages which a forecast price change must exceed before a commitment should be profitable. This is, therefore, the general criterion for commitments, either for buying or selling for all forecasting intervals.

The specific criteria for buying and selling are obtained by the application of these general criteria to the various interval probability curves. For instance, when the stock is at 100, the required price change to cover commitment costs is 0.7 per cent. With the 8-8 interval, this change is practicable when the stock during the prior interval has changed 1.4 per cent. Thus, at that price level 1.5 per cent is the critical point for making a commitment.

The practical application of this method in actual tests will now be considered. The hypothetical market trading here carried through is a comparatively rapid, inexpensive, yet a correct way in which to test trading theories.

The procedure is simple and will be illustrated in some detail by the test of the 1-1 interval criteria. Reference has already been made to the

daily percentage changes in the Steel Common. These percentages were originally tabulated on the work sheet in a column adjacent to the stock price. Beginning with January 1, 1922, the eye runs down this column until January 20, when the price change of plus 2.63 per cent exceeded 2.00 per cent, which at the market level was required to meet the round turn cost of the transaction. The stock was assumed to have been bought as of that date at 88-1/8. It was not until June 19 that a minus change of more than 2.69 per cent occurred, the change amounting to 2.40 per cent. The stock was assumed sold at 99-2/8, the "high" for that day, and a short position taken at that same price. On October 8 a positive change of 2.12 per cent occurred and the market position was reversed, this time at the "high" for the day of 102-2/8. This position was maintained until the end of the year when the price was 107-6/8.

Thus, for this year's trading the results were a gain of 11-2/8 points, a loss of 9-2/8 points with the stock held at further loss of 0-4/8 points, a gross profit of 1-4/8 points. The costs of these transactions were 32, 70, and 72 cents, a total of $1.74, giving the net result for the year of $0.24 loss. This includes the brokerage which would have been charged if the account had been closed.

It was observed that for the ten year period there was a gross profit of 189 points, a trading cost of 90.33, leaving a net profit from the 125 transactions of 92.67. That averages only $9.00 per year per share. The rate of profit was quite meager and the year to year record was quite erratic. Actual net losses were sustained in four of the ten years, and in two years amounted to as high as $45.00 per share.

While the hypothetical market trading demonstrates the theory, it also proves the 1-1 criterion was anything but a practical one to employ in trading. Hypothetical accounts were determined in the same manner for the other intervals, whose probability curves are shown in Chart I, the summaries of which are shown in Table II. The gauges of practicability, of course, are the items of net profit and the consistency of the profits throughout the decade.

Note well the trend of total net profits in the Decade Summary. As the criterion intervals are increased to 16-16, the total net gains increase to a maximum, after which they decrease. This is in harmony with the indications of the Summary Probability Characteristics in Table I. Moreover, as the same intervals were increased, the total of the annual net losses was decreased until the 32-32 criterion interval yielded a minimum. It must be concluded from these facts that the most satisfactory trading criterion will be in the region of the 16-16 and 32-32 intervals.

This conclusion prompted a further investigation of the trading possibilities of these two criteria in confirmation. The procedure was but slightly different than with the single criterion. The first commitment was made when both criteria, concurrently, indicated that the stock should be bought or sold. The position was maintained until both criteria, concurrently, indicated that the position should be reversed.

TABLE II

DECADE SUMMARIES OF HYPOTHETICAL TRADING

Intervals	Total annual net losses in points	Total annual net gains in points	Total net gains in points	Total number of commitment reversals
1-1	92.07	187.32	92.07	125
2-2	85.17	247.09	162.40	174
4-4	41.14	317.39	276.25	169
8-8	23.44	295.34	271.90	222
16-16	33.82	382.49	348.67	183
32-32	11.36	297.08	280.50	110
48-32	92.67	173.11	110.54	157
64-32	128.00	120.05	0.00	87
16-16 with 32-32 in confirmation	15.63	251.12	235.49	95

Table II shows that with this criterion, the total net gains were decreased from the maximum of 348.67 points to 235.49 points, the total annual net losses were also decreased from the minimum of 33.82 points to 15.63 points. While these tests are too crude to be the basis of exact judgments, the writer feels that these confirming criteria must approximate the most satisfactory one that it is possible to use as a basis for market trading in U. S. Steel Common stock.

From a statistical standpoint, the results obtained in the foregoing are not fully conclusive. The analysis has the inherent fault that the test covers the same period from which the curves themselves were derived.

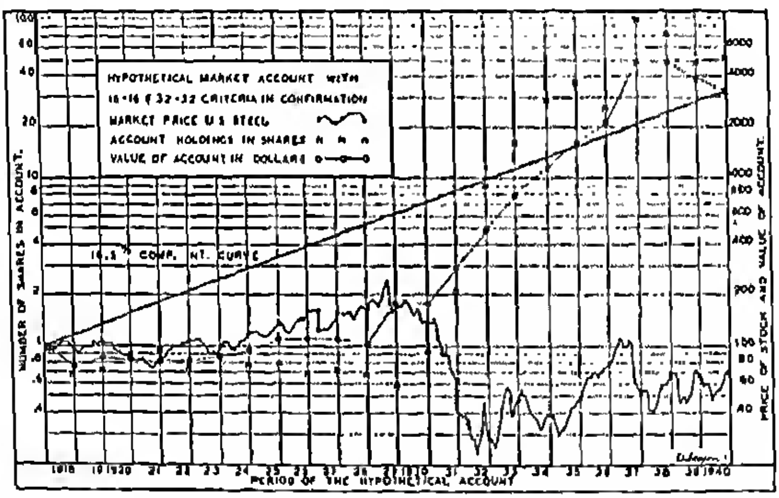
This is entirely satisfactory for the short interval forecasting curves, such as the 2-2, 4-4 or perhaps even the 8-8 intervals. It is quite possible, however, that with the longer intervals, accidental cycles have given a character to the 16-16 and 32-32 interval probability curves that would not be duplicated in another period.

For this reason, a more detailed investigation has been made of the forecasting merits of these two criteria in confirmation. The period for which hypothetical market operations were executed includes the years 1918 to 1921, the originally analyzed period of 1922 to 1931, and also the years 1932 to 1940, each inclusive, the three periods of four, ten and nine years each, making a total of twenty-three years. This long

period should exclude the possibility of accidental successes and yield results in which great confidence may be placed.

While the commitment procedure was the same as previously employed, the accounting procedure was more detailed. The amount of trading account was computed at the completion of each commitment. Furthermore, in the following commitment, the full amount of the trading account is assumed to be employed, all stock bought outright, or if the commitment was short, the amount sold was given a coverage equal to the price of the stock at the time of the sale. In both cases, for accurate accounting, this necessitated computing the extent of the commitments to a ten thousandth of a share.

CHART III



The account was opened with $100.00 and the first commitment was on January 8, when the stock was bought at a price of 95-7/8. The round turn brokerage and taxes amounted to $64.00 for 100 shares so that 1.0361 shares were purchased at a per share cost of $96.51. The stock was sold March 5 at 91-2/8 for a loss of 4-5/8 points and a net loss, including brokerage and taxes on the 1.0361 shares, of $5.46. The account then stood at $94.54. This full amount was employed in the short sale on the same day at 91-2/8, which after deducting for brokerage and taxes permitted a sale of 1.0282 shares. The commitment was covered April 18 at 94-5/8 for a second loss of 3-3/8 points, and a net loss, including brokerage and taxes on the 1.0282 shares of $4.1900. This brought the new balance down to $90.35.

Such was the procedure followed for the twenty-three years. The only departure was the rendering of balances at the end of each year for charting purposes. In Chart III are shown the records of the main items of the account, the stock prices, the extent of the holdings and the amounts of the account.

Chart III warrants close examination for it summarizes the entire study. The first year a loss of 23.85 per cent was incurred, while a period of eight years elapsed before the loss was finally recouped. For three more years the results were practically neutral. During the next decade there was a spurt in which the account was multiplied nearly fifty fold, only to be reduced by a third in the next and final two years.

During the ten year period 1922 to 1931, inclusive, the account increased from $84.45 to $171.16, an increase of but 7.4 per cent compounded annually. This is a clear indication that the period from which the criterion was established had little to do with the success over the entire 23 year period. Over this period, the account increased from $100.00 to $3368.77, or at the rate of 16.5 per cent, compounded annually.

This entire analysis permits making four important conclusions.

1. The interval probability method of analysis of market action demonstrates there are no well defined cycles in U. S. Steel Common stock, or in the wheat and cotton futures market.

2. Forecasts of U. S. Steel Common stock based on the probability curves of intervals up to the 32-32 day intervals all yield net profits in various amounts.

3. Maximum profits are obtainable when the forecasts are based upon the action of the market over an interval of about 32 market days.

4. Even the best interval probability forecasting results are so erratic from year to year, that the usefulness of this method of market forecasting is seriously impaired.

THE USE OF INVERSIONS AS A TEST OF RANDOM ORDER

By A. C. Rosander
War Production Board

It can be shown that the inversions in position of n objects or magnitudes 1, 2, 3, . . . , n for the $n!$ possible permutations are distributed in a family of symmetrical frequency distributions with moments which are functions of n only. In counting inversions of position the natural order 1, 2, 3, . . . , n is used as the criterion. All possible inversions for two and for three objects are as follows:

2 objects		3 objects	
Permutation	Number of inversions	Permutation	Number of inversions
1 2	0	1 2 3	0
2 1	1	1 3 2	1
		2 1 3	1
		2 3 1	2
		3 1 2	2
		2 1	3

These lead to the following frequency distributions which can be extended to include n objects or magnitudes:

Number of inversions	$n=2$ Frequency	$n=3$ Frequency
0	1	1
1	1	2
2		2
3		1
Sum ($n!$)	2	6

The general rule to follow in counting inversions is to take each rank and count how many lower ranks follow it; the sum of all such counts is the number of inversions for that permutation.

In the inversion distribution the expected number of inversions per permutation, the arithmetic mean, is

$$M = \frac{n(n-1)}{4}$$

while the variance is

$$\sigma^2 = \frac{n(n-1)(2n+5)}{72}.$$

The total number of inversions is $n!\ n(n-1)/4$ while the area of the frequency histogram is $n!$.

It can be shown further that this symmetrical single-peaked distribution approaches quite rapidly the normal curve as n increases indefinitely.[1] Even when n is 6 the correspondence is quite close. In this respect and in several others the distribution is analogous to the point binomial. This means that no large error will be committed if we assume that the arithmetic means from a large number of random samples are distributed as a normal probability curve with a mean of M and a variance of σ^2/N where N is the size of the sample and σ^2 is the population variance indicated above.

These distributions give us a pattern which can be used to test some hypothesis of order. The hypothesis provides the basis of counting inversions, and the frequency distribution allows us to make inferences relative to the departure of the data from the hypothesis. Hence these inversions distributions may be taken to give an operational definition of random order; deviates of these distributions are measures of the departure of the data from random order.

As an exploratory technique, the method of inversions appears to have some merit. Its use, however, is based upon ranked data which represent in general a loss of information. It assumes that the various permutations of order in the data are equally likely. It does not allow for tied ranks. Furthermore it is imperative that the hypothesis of order to be tested is independent of the data. Even with limitations there is a number of problems to which these distributions are applicable.

APPLICATIONS

The Randomness of Tippett's Numbers. Ten sets of Tippett's four-digit numbers were selected, in groups of 25, from the first page of his table of random numbers. The expected number of inversions per permutation for 25 orders is 150, the variance is 458, and the standard deviation is 21.4. The results are as follows:

[1] See M. G. Kendall, *Biometrika*, June 1938; and George B. Dantzig, *Annals of Mathematical Statistics*, September 1939. The writer first developed and used the method of inversions in 1937 independent of these two investigators.

Set number	Mean $\bar{x}$	$\frac{\bar{x}-M}{\sigma}$	Mean of cumulated sets	Deviation of cumulative mean from $M=150$
1	101	−2.3	101	−49
2	174	1.2	138	−12
3	173	1.1	149	− 1
4	170	0.9	155	5
5	132	−0.9	150	0
6	168	0.9	153	3
7	125	−1.2	149	− 1
8	139	−0.5	148	− 2
9	179	1.4	151	1
10	177	1.3	154	4

The mean of the 10 sets or samples is 153.8; since the expected mean is 150, the deviation is 3.8. Is this deviation to be expected on the basis of sampling fluctuations? The standard deviation of these means based on a sample of 10 is

$$\sigma_{\bar{x}} = \frac{21.4}{\sqrt{10}} = 6.77.$$

On this basis the deviation of 3.8 is 3.8/6.8 or 0.56 standard deviations above the expected mean. Assuming that the means are distributed approximately as a normal probability curve for samples of 10, such deviations have a high frequency of occurrence. Hence we infer that this sample of 10 could represent a population of random numbers. As additional evidence we notice that no one of the ten inversion values deviates as much as 2.5 standard deviations from the mean of 150; four deviates are negative while six are positive. All this is evidence that according to the criterion of inversions these groups of 25 four-digit numbers represent random arrangements.

The Randomness of Measurements in Quality Control. Two hundred and four measurements given by Shewhart[2] were tested in groups of 51 and also in a single group of 204 ranks. Since there were quite a number of tied ranks, these were broken by assigning the ranks involved at random to the values in question.

For the four groups of 51 ranks the expected number of inversions is 637.5 while the standard deviation is 61.6. The results are as follows:

[2] Walter A. Shewhart and W. Edwards Deming (editor), *Statistical Method from the Viewpoint of Quality Control*, Washington, D. C. The Graduate School of the Department of Agriculture, 1939, pp. 32, 90.

Set	Mean $\bar{x}$	Expected number or Mean M	$\bar{x}-M$	$\frac{\bar{x}-M}{\sigma}$
1	539	637.5	− 98.5	−1.60
2	528	637.5	−109.5	−1.78
3	603	637.5	− 34.5	−0.56
4	712	637.5	74.5	1.21

The greatest deviation from the expected number is 1.78 standard deviations which is not an uncommon deviation to find. There seems no particular reason to reject any of these sets as representing a non-random order, on the basis of our criterion of inversions.

When we come to test the 204 measurements as a single sample we find a different situation. For 204 ranks the expected number of inversions is 10,353, while the population standard deviation is 489. By actual count the total number of inversions is 8,054 or 1,399 below the expected value. In terms of standard deviations this is −2.86 units, a deviation which occurs only once in about 500 times based on the normal probability distribution. Examination of the measurements indicates why the number of inversions falls so far below the expected number based upon random order. From observation 17 to observation 148 inclusive there is a downward trend of the values; from observation 149 to the end all measurements are at a high level; only 8 out of 56 measurements fell below 4,500. Above the 148th measurement we find 42 measurements which rank above 102, the mid-rank, and only 14 rank below. This preponderance of large measurements at the upper end of the series accounts for the relatively low number of inversions.

These tests indicate an interesting point: that within a non-random series there may be random sets, while within random series there may be non-random sets. Certainly whether one obtains a test favorable to the hypothesis of randomness, on the basis of this criterion, depends on where one starts in the series and where one ends. Presumably if one had not 204 measurements but many times this number, providing that there was no better statistical control, he would obtain no better indication of random order.

What this indicates is that one may obtain a test of random order by the mere connecting together of two or more sets of non-random sequences, and vice versa. Another point those in sociology and economics will observe is that this sequence of measurements in the electrical laboratory appears to be strikingly similar to the order of the magnitudes so commonly found in economic time series where any type of statistical control is out of the question.

The Randomness of Economic Time Series. In the data from Shewhart we found that the parts gave a test favorable to random order, whereas a test of the total gave a test favorable to the hypothesis of non-random order. We turn now to an example from an economic time series where just the opposite is the case. The consumption of sulphuric acid by various industries in the United States during the past 13 years takes the following ranks:

1928	7	1932	13	1936	6
1929	3	1933	12	1937	2
1930	5	1934	10	1938	9
1931	11	1935	8	1939	4
				1940	1

For 13 ranks the expected number of inversions is 39, the standard deviation is 8.2. In the above table the total number of inversions is 50 or 11 above the mean, or about 1.4 standard deviations above the mean. This is not an unusual deviation to find; hence one may on the basis of this criterion consider this a random order. An examination of the last 9 years gives different results. For this number of ranks the expected number of inversions is 18; the standard deviation is 4.8. From the data we obtain 32 inversions, a deviation of 14 or about 2.9 standard deviations. One is tempted with such a deviation to be skeptical of the randomness of the measurements involved.

Furthermore we might test the hypothesis that the order of the values from 1932 through 1940 was the same as the order of the Federal Reserve Board index of industrial production (1925–1939 = 100). The order of corresponding years of the two sets of data is the same; the number of inversions is therefore zero. This is additional evidence which calls in question the hypothesis of randomness.

The Randomness of Guessing. The probability of selecting by chance the correct order of n magnitudes is $1/n!$. In an examination question in which 3 elements are to be arranged in some sequence, the probability of getting the correct sequence on the basis of complete ignorance is 1 in 6 which is quite high. On the other hand, if the number of elements is increased to 5, the chance of guessing a correct sequence is reduced to 1 in 120.

This principle could be used in the extra-sensory perception experiments of Rhino's in which the numbers of inversions in the order in which the first n numbers are presented could be used as a measure of departure from chance selection. By using a sequence of 10 magnitudes the probability of selecting by chance the correct order is reduced to 1 in 3,628,800.

The method of inversions can be extended to provide an operational basis of measuring discrimination of a sequence of objects or measurements. On the basis of this theory the expected number of inversions is a measure of no discrimination, or guessing, and departure from the expected number is a measure of ability to discriminate. Hence it is possible to give a "score" of discrimination to each of the $n!$ possible permutations.

Random Order and Correlation.[a] In case two sequences are uncorrelated, the inversions of order of one relative to another will be distributed according to the foregoing distributions. In other words, zero correlation is associated with the expected number of inversions, M.

On the other hand, if two sequences are correlated the number of inversions will tend to be small for a high positive correlation, and to be large for a high negative correlation.

These relationships suggest the following definition of a correlation function r_i in terms of the number of inversions x:

$$r_i = 1 - \frac{2x}{x_m} = 1 - \frac{4x}{n(n-1)}$$

where x_m is a constant, the maximum number of inversions for a given n.

When x, the expected number of inversions, is equal to zero, $r_i = 1$; when it is equal to M which is always one-half x_m, $r_i = 0$; when it is equal to x_m, $r_i = -1$.

The value of this simple linear function lies in the fact that the distribution of x, for any given value of n, is known and can be used to evaluate any value of r_i. Hence there is no need of dealing with the distribution of the correlation coefficient. For values of n larger than 10 one can make use of the normal probability curve as an approximation to the inversion frequency distribution, and test whether the given value of x falls within the 5 per cent or 1 per cent levels.

Suppose we have two sets of ranks, A and B, with the set A ranked from 1 to 20, with the corresponding rank of set B below that of set A, as follows:

Set A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Set B	11	6	7	2	3	10	1	4	12	8	5	9	14	18	19	15	17	13	16	20
x_i	10	5	5	1	1	5	0	0	3	1	0	0	1	4	4	1	2	0	0	0

[a] This method appears to have been published first by M. G. Kendall, *Biometrika*, June 1938, although the writer arrived at a similar index and similar conclusions independently. The writer prefers to use inversions and their distributions rather than correlation because the former can be associated with testing some hypothesis of order which covers a wider range of problems than does the correlation coefficient. Kendall uses the standard error of r but this is unnecessary since r is simply a linear function of the number of inversions.

Once the two sets of ranks are arranged in this manner, the next step is to count the total number of inversions. This is done by beginning at the left end of the B series, taking each number in turn, and counting the number of ranks to the right which are smaller than the one under consideration. The numbers of inversions corresponding to the individual ranks of set B are shown in the row designated "x_i." The total number of inversions is the sum of this row, or 43. It does not make any difference whether we use set A as the standard order to count the inversions in set B, or vice versa; the number of inversions is the same.

When n is 20, the expected number of inversions is 95, the 5 per cent point is 70, and the 1 per cent point is 59. In other words, a value less than the obtained value of 43 could be obtained by chance less than once in 100 times. On the basis of this criterion we would reject the hypothesis that there is no relation between these two ranks. Then if we desire to translate this inferred relation into an index of correlation we can substitute the x value of 43 in the equation for r_t. If we do this we obtain a value of 0.55 which is positive because x is less than the mean; it would be negative if x were greater than the mean.

CORRELATION ANALYSIS BY MARGINS

By E. J. Broster*

Correlation analysis by the use of margins (or differences) is not entirely a new idea except in so far as little seems to have been done to exploit it, or, anyway, to demonstrate its potential value as a statistical device.[1] Like any method—or may I say any *other* method—of correlation analysis worthy of the name, the use of margins has special value in the determination and even the final solution of at least one particular functional type of problem. In short, it fills a gap. But it can also be successfully applied to other kinds of problems, as I know from experience—and with advantage too, where the method of least squares is beyond the mathematical understanding of those who may have to make use of the resulting coefficients.

The particular gap which the marginal method fills concerns the solution of problems in one independent variable in which graphing indicates the existence of an equation of some unknown but higher degree than the first, that is, of an equation of the type:

$$Y=k+aX+bX^2+cX^3+\cdots$$

By using margins, we can obtain the correct or best-fitting degree of equation for final solution by least squares, or we can continue the analysis by margins to arrive at reasonably accurate results.

Having fitted a representative curve to the coordinates either freehand or on the basis of group averages, we proceed by recording a number of readings from it. For a reason that will become clear later, these should be taken in order of magnitude of the independent variable, X. In Table I, lines (a) and (b) give the readings of a simple but typical example. The rest of the table down to line (l) shows the method of extracting margins up to the point where the degree of the equation is determined together with the coefficient, or an approximation to it, of X^n, where n is the degree of the equation.

In this case $n=3$. We know this because Y/X of the third series of margins in line (l) are the same, or at least the trend shown in the first

* The following quotation from the author's letter of transmittal should be of interest to many of us. "Most of the preliminary experimental work on the use of margins I carried out in my leisure evenings during the heavy air raids on the London area last year when I found it difficult to concentrate on the complexities of the kind of investigations I specialise in." Editor

[1] See H. S. Will, "On Fitting Curves to Observational Series by the Method of Differences," *Annals of Mathematical Statistics*, May 1930; S. S. Bose, "Relative Efficiency of Regression Coefficients Estimated by the Method of Finite Differences," *Sankhyā: The Indian Journal of Statistics*, Part 4, 1938; and M. Ezekiel, *Methods of Correlation Analysis*, Table 13, p. 46.

TABLE 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Determining the degree of equation and the coefficient of X^n						
Readings from graph:—							
(a)	X	1	2	3	4	5	5.5
(b)	Y	0	27	71	153	285	373.5
Margins—First series, to eliminate k:—							
(c)	X'	—	1	1	1	1	0.5
(d)	Y'	—	18	44	82	132	88.5
(e)	Y'/X'	—	18	44	82	132	177
	Second series, to eliminate aX:—						
(f)	X''	—	—	2	2	2	1.5
(g)	Y''	—	—	26	38	50	45
(h)	Y''/X''	—	—	13	19	25	30
	Third series, to eliminate bX^2:—						
(i)	X'''	—	—	—	3	3	2.5
(k)	Y'''	—	—	—	6	6	5
(l)	Y'''/X'''	—	—	—	2	2	2
	The equation is therefore of the third degree and the coefficient of X^3 is 2.						
	Determining the coefficient of X^{n-1} (i.e., X^2)						
(m)	$Y-2X^3$	7	11	17	25	35	40.75
Margins—First series, to eliminate k:—							
(n)	X'	—	1	1	1	1	0.5
(o)	Y_1'	—	4	6	8	10	5.75
(p)	Y_1'/X'	—	4	6	8	10	11.5
	Second series, to eliminate aX:—						
(q)	X''	—	—	2	2	2	1.5
(r)	Y_1''	—	—	2	2	2	1.5
(s)	Y_1''/X''	—	—	1	1	1	1
	The coefficient of X^2 is therefore 1.						

and second margins, lines (c) and (h), is eliminated. In practice, identical values for any Y/X series is unusual, so that the criterion is elimination of trend, which indicates the need for setting out the original readings in order of magnitude of X.

The calculation of the first marginal series is clear. Line (c) is inserted to show the need for calculating the Y'/X' figures. It will not only rarely be found possible to arrange for each of the X' margins to equal unity but even be advisable to set out the X readings so that the first series of margins increase or decrease progressively. The reason for this is that as a result of graphic errors, the need for rounding off in the calculations and the accumulating error these involve with each successive extraction of margins, the final trendless Y/X series—when it is reached—will sometimes scarcely be recognized as such owing to the fluctuations it contains. When the readings taken from the graph result in variations in the X margins, these margins can be plotted against the corresponding Y margins. If they do give a trendless Y/X series, then the coordinates must necessarily lie about a straight line which passes through the origin. The slope of this line would be equal

to the coefficient of X^n, a fact which it is necessary to keep in mind if it is proposed to complete the analysis by the marginal method.

After the first extraction, the difference in the means of arriving at further margins should be noted. The Y margins are simply derived from the preceding Y/X series, while each figure in the X marginal series is equal to the difference between the two extreme figures involved in the original X readings, a difference which equals the sum of the corresponding figures in the first series of X margins. In Table I for instance, the figure of 2.5 in column (7), line (i), is equal to $5.5-3$ in line (a), and to $1+1+0.5$ in line (c). It is generally more convenient in practice to use the latter, and if necessary to apply the former as a check on the calculations.[2]

It is clear from the small number of marginal values left over at the end that a larger number of readings than those used in the example are necessary. The minimum is about ten, but a great deal depends upon the dispersion of the coordinates in the original diagram and the degree of rounding off necessary in the calculations. Dispersion always gives rise to difficulties of curve fitting, but even supposing the graph approximates closely to the type under consideration, the larger the number of readings taken for the purpose of the analysis the more accurate are the results likely to be.

If we propose completing the analysis by margins without resorting to least squares then the next step is to determine the value of b by deducting cX^3 from the Y series and repeating the process. This is shown in lines (m) to (s) of Table I, where it is found that $b=1$. Further analysis on the same lines would produce the complete equation:

$$Y = 5 + X + X^2 + 2X^3.$$

[2] Each final Y/X margin is in effect the solution for the coefficient of X^n, by simultaneous equations, of the sets of observations involved. For the second degree equation

$$Y = k + aX + bX^2$$

the solution for b from the three first sets of observations X_0, Y_0; X_1, Y_1; X_2, Y_2; by simultaneous equations, may be written

$$b = \frac{\dfrac{X_2 - X_1}{X_1 - X_0}(Y_1 - Y_0) - (Y_2 - Y_1)}{\dfrac{X_2 - X_1}{X_1 - X_0}(X_1^2 - X_0^2) - (X_2^2 - X_1^2)}.$$

This is identical to the solution obtained by the method of margins, viz.:

$$b = \frac{Y''}{X''} = \frac{\dfrac{Y_2 - Y_1}{X_2 - X_1} - \dfrac{Y_1 - Y_0}{X_1 - X_0}}{X_2 - X_0}.$$

We can regard each solution for b as an estimate. From the three first sets of observations, let this estimate be b_0. Then we derive b_1 from X_1, Y_1; X_2, Y_2; X_3, Y_3; b_2 from X_2, Y_2; X_3, Y_3; X_4, Y_4; and so on. Perfectly correlated data would give $b_0 = b_1 = b_2 \ldots$

This kind of marginal analysis may be applied to any curve in any quadrant, so that the constants k, a, b, c, . . . can be integral or fractional, positive or negative numbers. The indices are always necessarily positive integers. In the example all readings and margins have positive signs. Negative signs however often appear at some stage of an analysis, when the ordinary rules apply. Every analysis indicates the sign of each constant. Table II shows the first stage of an analysis involving three quadrants, since the complete equation is

$$Y = -10 + X + 3X^2 - X^3.$$

It should be kept in mind that the correct form of equation stating the relationship between any two variables may be of an entirely different nature. It is not surprising therefore that marginal analysis sometimes fails to yield results. Where this is so, its inexpedience or that of the manner of its application soon becomes evident in an increasing, instead of a decreasing, upward or downward trend in the successive Y/X series, or in a violent reversal of trend.

TABLE II

Determining the degree of equation and the coefficient of X^n

Readings from graph:—							
X	−3	−2	−1	0	1	2	3
Y	41	8	−7	−10	−7	−4	−7
Margins—First series:—							
X'	—	1	1	1	1	1	1
$Y' = Y'/X'$	—	−33	−15	−3	3	3	−3
Second series:—							
X''	—	—	2	2	2	2	2
Y''	—	—	18	12	6	0	−6
Y''/X''	—	—	9	6	3	0	−3
Third series:—							
X'''	—	—	—	3	3	3	3
Y'''	—	—	—	−3	−3	−3	−3
Y'''/X'''	—	—	—	−1	−1	−1	−1

The equation is therefore of the third degree, the coefficient of X^3 being 1 and its sign negative.

On the other hand, a good fit by marginal analysis does not provide any guarantee that the equation obtained states the true relationship. This method, in common with all other methods of correlation analysis, invariably gives empirical formulae except possibly where the correct form of equation is known and used.

A reversal of trend in the Y/X series does not necessarily condemn the method as being unsuitable for the particular case. It should generally be regarded as a signal that the first stage of the analysis has been reached. It is then a matter of choosing between the pre-reversal and the first reversal series, and the X and Y series corresponding to the

one with the smaller trend should be plotted in the way described above in order that a final decision can be made.

Where a violent reversal of trend takes place, the analysis should be discarded, but not necessarily the method. A good fit may sometimes be obtained on the assumption that the best-fitting equation involves negative indices which need to be converted to positive indices. For instance, it may happen that a good fit can be obtained on the basis of

$$Y = a\frac{1}{X} + k + bX + \cdots$$

which, converted for marginal analysis, becomes

$$YX = a + kX + bX^2 + \cdots$$

For the purpose of the calculations, YX would be regarded as the dependent variable, and treated in exactly the same way as Y in the example above.

Another useful application of the marginal method lies in the partial solution of problems in two variables, especially where the number of observations is too small, or the nature of the dependent variable's relationship to one of the two independent variables is too uncertain, to permit of the use of least squares without preliminary tests.

Table III contains some figures from the Board of Trade's Production Index. Columns (1), (2) and (3) give respectively the years, the general index numbers and the gas and electricity index numbers. The output of gas and electricity depends principally upon industrial production as a whole; but there are also some other factors which are responsible for the obvious upward trend in time independent of industrial production, and which can therefore be represented by the catch-all factor, *time*, statistically represented by the series of years.

TABLE III

Year		Board of Trade Production Indices*		Marginal series				
		General	Gas and electricity					
	t	X	Y	t'	X'	Y'	$67.6t$	$Y - 67.6t$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		(1924 = 1000)						
1927	0	1,068	1,107	—	—	—	0	1,107
1928	1	1,055	1,200	1	−13	93	68	1,132
1929	2	1,118	1,358	1	63	98	135	1,223
1930	3	1,032	1,387	1	−86	29	203	1,184
1931	4	937	1,424	1	−95	37	270	1,154
1932	5	933	1,470	1	−4	46	338	1,132
1933	6	986	1,562	1	53	92	406	1,156
1934	7	1,108	1,700	1	122	138	473	1,227

* Source: *Statistical Abstract for the United Kingdom*, 1940, p. 307.

the groups, the two second, and so on, and divides the sum of the Y margins by the sum of the X margins) merely provides a roundabout way of determining the regression coefficient from two-group averages; for the sum of the half-range margins in each series is equal to the difference between the sums of the basic observations in the two groups.[4]

For simple linear correlation, the method of margins is superfluous. It has no advantages even over the graphic and group-average methods, and has one or two serious disadvantages. As I hope I have shown, however, if it is employed with care and with the aid of graphs to avoid summation, it undoubtedly provides a simple means, in some of the more complex problems, of determining empirical forms of basic equation where pure logic fails, as it so often does in the case of economic and financial relationships, to give the correct forms.

[4] H. S. Will (loc. cit., pp. 105–106) like Rose, appears to have overlooked this fact. His formula for linear series makes provision for any number of steps, k, which, if it happens to be a factor of the number of observations, n, gives the same solution as n/k-group averages except where $k = 1$, when the solution is based entirely upon the two extreme sets of observations. Will also provides formulae for calculating the regression coefficients of other types of equation by the method of margins, but their application is restricted to problems in which the form of equation is logically determinable.

THE STANDARD ERROR OF PERCENTILES

By W. Duane Evans
U. S. Bureau of Labor Statistics

In recent studies, percentiles of various kinds have been much used to demonstrate economic differences between segments of populations. This use has been specially marked in surveys of the distribution of persons or families according to their incomes.

Percentiles have much to recommend them. They have a specific meaning in terms of population values, and they may be defined in terms which are readily understood by a person without technical training. Moreover, they are in general somewhat less sensitive to errors resulting from certain types of sampling bias than, for example, the mean.[1]

One of the drawbacks to the use of percentiles has been the lack of some general method by which the probable error due to chance deviations in sampling might be investigated. With respect to the median, the rule that the standard error is about one and one-quarter times the easily estimated standard error of the mean is well known. However, this rule is applicable only when the population distribution approximates the normal form. It is not valid for the highly skewed and flattened distributions in which incomes usually fall. In fact, in such cases the standard error of the median is generally substantially less than that of the mean.

The following paragraphs present a method by which the standard error of any percentile may be estimated. In the case of large samples this requires in general less effort and computation than an estimate of the standard error of the mean. For large samples, the method has been worked out for the general case of any percentile of a sub-group of a population estimated by means of a stratified sample. The results may, of course, be readily simplified to apply to less complex systems of sampling. The method does not require the assumption that the parent population is distributed in any special way.

The method presented does not result in exact values for the standard error of percentiles, except in the limit, but rather in usable and in many cases very close approximations to these values. The principal

[1] It has been noted that high income families tend to be somewhat less cooperative in income and expenditure studies than families in the middle ranges. Even though high income families are infrequent in the population, a slight bias against them, resulting in a disproportionately low number of such families in a sample, may markedly change the position of the estimated mean while hardly altering the estimate of the median.

aim is to provide the practical statistical worker with a simple means of testing the significance of percentiles and the equipment necessary to design surveys with preassigned levels of reliability.

Large Samples. The population is defined as consisting of N sampling units or individuals. Among these are A sampling units possessing some complex of characters which sets them apart from the rest. The A type individuals vary with respect to some additional factor X. This factor is of such type and the number A is sufficiently large so that the variation is effectively continuous. Corresponding to a preassigned value of the character X, say X_B, there are among the A type individuals a certain number B, for which the value of X is less than X_B.

To assist in visualizing the foregoing, we may, for example, let N represent the number of individuals in a city. The symbol A may then represent the number of persons within the city that are of some specific race and sex. These vary according to their incomes (X). There are then a total of B individuals of the specified race and sex which have incomes of less than, say \$1,000 (X_B).

The population is divided into r independent strata (to carry out the analogy of the city, into r distinct geographical districts). We then let—

N_i = the number of sampling units in the i^{th} stratum,

A_i = the number of A type sampling units in the i^{th} stratum,

B_i = the number of B type sampling units in the i^{th} stratum.

A sample is selected at random and without replacement in each of the strata. With respect to this sample, let—

n_i = the number of sampling units constituting the sample in the i^{th} stratum,

a_i = the number of A type sampling units in the sample n_i,

b_i = the number of B type sampling units in the sample n_i.

Let the reciprocal of the sampling ratio employed in the i^{th} stratum be represented by S_i. Then

$$S_i = N_i/n_i. \tag{1}$$

Estimates of quantities are denoted by primes. Estimates of A and B are then defined by the relations

$$A' = \sum_{i=1}^{r} S_i a_i \tag{2a}$$

$$B' = \sum_{i=1}^{r} S_i b_i. \tag{2b}$$

It is evident that $E(A') = A$ and $E(B') = B$.

A percentile, X_K, is to be determined for the group of A type individuals in the population. Let the percentile be defined by K, which represents the proportion of the A type individuals to which the percentile is to refer. Then, for example, $K=0.25$ for the first quartile, or $K=0.50$ for the median.

The quantity Δ is defined by the equation

$$\Delta = KA - B. \tag{3}$$

It then represents the difference between the number of individuals of type A in the population with a value of the character X less than X_K and the number with a value of the character X less than X_B. An estimate of Δ is obtained as follows:

$$\Delta' = \sum_{i=1}^{r} S_i(Ka_i - b_i). \tag{4}$$

It may be shown that

$$E(\Delta') = \sum_{i=1}^{r} S_i n_i(Kp_{a_i} - p_{b_i}) = \Delta \tag{5}$$

where p_{a_i} is the proportion or probability of occurence in the i^{th} stratum of individuals of the A type. The definition of p_{b_i} is similar, and both may be represented as follows:

$$p_{a_i} = A_i/N_i \qquad\qquad p_{b_i} = B_i/N_i.$$

We may now note a very important fact. On the basis of any particular sample, we may estimate the percentile X_K to be either above or below the value X_B. However, the latter can occur *only* if the frequency difference Δ' for this particular sample is negative in sign. Then, if we can determine the relative frequency of occurence or probability of negative values of Δ' under any particular set of conditions, we at the same time have specified the relative proportion of all times when on repeated sampling we will estimate the percentile X_K to lie below X_B. This may be expressed by the relationship

$$p(\Delta' < 0) = p(X'_K < X_B) = \int_{-\infty}^{0} f(\Delta')d\Delta' \tag{6}$$

where $f(\Delta')$ represents the probability distribution of the estimated frequency differences. This relationship is true irrespective of either population or sample size.

We may now proceed to examine the form of the distribution of Δ' in samples of large absolute size. Since the quantity Δ' is a linear func-

tion of a number of hypergeometrically distributed variates, one may at once infer that as the size of the sample on which it is based increases while the ratio of sample to population size remains small, the distribution of Δ' will approach the normal probability distribution as a limit. The latter form will give a fair approximation to the probability distribution of Δ' in the case of samples of even moderate size drawn from a large population.

Under the given conditions, equation (6) may then be rewritten as follows:

$$p(X'_K < X_H) = (1/\sqrt{2\pi}) \int_{-\infty}^{-t} e^{-z^2/2} dz \tag{7}$$

where

$$t = \Delta/\sigma_{\Delta'}. \tag{8}$$

In evaluating (7), it is necessary to have the variance of the distribution of Δ' in terms of the population parameters. This is easily obtained in the following manner. Since the individual strata are independent, it will be sufficient to determine the variance of Δ_i', where

$$\Delta_i' = S_i(Ka_i - b_i) \tag{9}$$

because, from (4),

$$\sigma^2_{\Delta'} = \sum_{i=1}^{r} \sigma^2_{\Delta_i'}. \tag{10}$$

From (9)

$$E(\Delta_i'^2) = S_i^2 E(K^2 a_i^2 + b_i^2 - 2Ka_i b_i). \tag{11}$$

But it may be shown that

$$E(a_i^2) = n_i^2 p_{a_i}^2 + n_i p_{a_i} q_{a_i}(N_i - n_i)/(N_i - 1) \tag{12a}$$

$$E(b_i^2) = n_i^2 p_{b_i}^2 + n_i p_{b_i} q_{b_i}(N_i - n_i)/(N_i - 1) \tag{12b}$$

$$E(a_i b_i) = n_i p_{a_i} p_{b_i} + n_i p_{b_i} q_{a_i}(N_i - n_i)/(N_i - 1). \tag{12c}$$

Since the variance of any variate is equal to the expected value of its square minus the square of its expected value, we may combine equations (5), (10), (11), and (12) as follows:

$$\sigma^2_{\Delta'} = \sum_{i=1}^{r} S_i^2 n_i [K(1 - K)p_{a_i} - (1 - 2K)(Kp_{a_i} - p_{b_i}) - (Kp_{a_i} - p_{b_i})^2][(N_i - n_i)/(N_i - 1)]. \tag{13}$$

If the appropriate population parameters are known or can be estimated, equation (7) can now be evaluated to determine the probability that a percentile may be estimated to lie below some specific value X_B. It is only necessary to evaluate t and refer to suitable tables for a normal probability distribution of unit standard deviation and zero mean. The desired probability will be represented by the area falling below minus t.

By repeating the procedure for two different values of X, say X_{B_1} and X_{B_2}, we may determine the mathematical probability that X_K will be estimated to lie between X_{B_1} and X_{B_2}. Choosing appropriate values of X, we may then reconstruct the probability distribution of X'_K over any desired range.

To obviate this awkward procedure, we must learn something of the form of distribution of X'_K. In the first place, interest centers only in the shape of the central portion of the distribution of X'_K; the extreme tails of the distribution are of little importance. This is equivalent to saying that equation (7) will be evaluated only for values of t which lead to probabilities of significant size, say for values between -3 and $+3$. Since the variance of Δ' decreases with an increase in the sample size, it is apparent from (8) that as the sample becomes larger the values which may be assigned to X which lead to values of t lying within any preassigned interval will cover a shorter and shorter range. In a large sample the entire range of such values will be small. Two important results follow from this fact.

First, the population distribution within this narrow range may be assumed to be approximately linear, since it was originally assumed that the distribution of the A type individuals with respect to X was continuous, though of unspecified form. As the size of sample increases and the range of relevant values of X becomes narrower, it follows that the relationship between Δ and these values of X approaches linearity.

Second, examine the form of equation (13). The same reasoning may be applied here. In this expression only the values of p_{bi} in the various strata depend on the value of X, and over a narrow range of values of X, the values of the p_{bi} will change but slightly. Over this short range, then, the variance of Δ' is effectively constant.

With $\sigma_{\Delta'}$ approaching constancy and Δ becoming a linear function of X, it is apparent from equation (8) that the relationship between t and X also approaches linearity in the range over which we are interested in determining the form of distribution of X'_K. But as this relationship becomes linear, it follows from (7) that the distribution of

X'_K necessarily approaches the normal form. In a sufficiently large sample, then, the distribution of estimates of any percentile is very closely approximated by a normal probability curve.

If we accept the distribution of X'_K as approximately normal, t is obviously equivalent to the number of standard deviations of X'_K represented by the absolute difference between X_K and X_B. We may therefore set down

$$\sigma_{X'_K} = \{|X_K - X_B|\,\sigma_{\Delta'}\}/\Delta. \tag{14}$$

This expression may be somewhat simplified. In the first place, equation (13) reduces directly to the following form:

$$\sigma^2_{\Delta'} = \sum_{i=1}^{r} S_i[K(1-K)A_i - (1-2K)\Delta_i - \Delta_i^2/N_i][(N_i - n_i)/(N_i - 1)]. \tag{15}$$

In line with our previous assumptions, the finite sample factor in the second pair of brackets may be taken as approximately equal to unity. Now A_i represents all individuals of the A type in the i^{th} stratum, while Δ_i represents only those of the A type in the i^{th} stratum who fall in the relatively narrow interval between X_K and X_B. It is apparent that in most cases Δ_i will be quite small relative to A_i, and the following expression will be a satisfactory approximation to the variance of Δ'.

$$\sigma^2_{\Delta'} = K(1-K)\sum_{i=1}^{r} S_i A_i. \tag{16}$$

Because of the modifying factors $K(1-K)$ and $(1-2K)$, the agreement between equations (15) and (16) will be least satisfactory in the case of the extreme percentiles. However, here again sample size is a factor. As the size of sample increases, the range of values of X for which an evaluation of (15) may be required will become narrower, and accordingly the maximum values of Δ_i which may appear will become smaller. The second and third terms within the brackets thus decrease in importance relative to the first.

Finally, we may assume that X is measured in relatively fine intervals, and choose X_B so that $X_K - X_B$ is equal to one-half C, where C is the width of the class interval by which the character X is measured and within which the frequencies of the A type individuals are tabulated. Then Δ will be very nearly equal to one-half F_c, where F_c is the total number of A type individuals in the population in the interval C which includes the percentile X_K. Making the indicated changes, and incorporating equation (16), equation (14) reduces to

$$\sigma_{X'_K} = C\left[K(1-K)\sum_{i=1}^{r} S_i A_i\right]^{1/2} \Big/ F_c. \tag{17}$$

This expression is an *approximation* to the standard error of a percentile. In applying it, then, due care must be exercised that the assumptions involved in its derivation are not violated. Perhaps most important, it should not be applied to extreme percentiles unless the sample is large enough to justify the assumption of normality in the text following equation (6), and the use of equation (16) as an approximation to equation (15).

In evaluating equation (17) in any particular instance it will probably be necessary to use sample values as approximations to the population parameters specified. It may well be noted, then, that the ratio C/F_c is in reality an approximation to the value of the reciprocal of the ordinate of the distribution of the A type individuals according to the character X at the point X_K. It follows that a considerable range of data may be used to estimate the most probable value of the ratio in any case where the sample results are somewhat irregular.

Equation (17) is quite general in form. It may be readily simplified to refer to less complex sampling situations. For example, let us assume that a percentile referring to the entire population is to be estimated on the basis of a non-stratified sample of size n. Equation (17) then reduces to

$$\sigma_{X'_K} = C[K(1-K)n]^{1/2}/f_c \tag{18}$$

where f_c represents the expected frequency of A type individuals in a sample of size n in the interval which includes the percentile. If the percentile to be studied is the median, equation (18) becomes simply

$$\sigma_{X'_{0.5}} = C\sqrt{n}/2f_c. \tag{19}$$

This will be recognized as the same formula as that given by Yule and Kendall, but derived by them by a somewhat different method.[2]

Small Samples. In the foregoing, the argument has been limited to large samples. In small samples the various distributions which have been studied may not be represented with sufficient accuracy by normal error functions, and consequently equation (7) and its simplifications become invalid. However, a line of attack on the problem of determining the probable error of an estimate of a percentile based on a small sample is indicated.

In the first place, since a small sample is seldom a stratified sample, we may limit our examination to non-stratified samples. Similarly to

[2] G. U. Yule and M. G. Kendall, *An Introduction to the Theory of Statistics*, London, 1937, p. 384.

the terminology previously used, let n, a, and b represent the sample size and the number of A and B type sample units in the sample. Now, provided that the proportions of the A and B type individuals in the population are known or can be estimated, we may estimate the probability of occurrence of combinations of particular values of a and b in repeated samples of size n. From these probabilities we may determine the probability of appearance of negative frequency differences and so the probability that the percentile X_K will be estimated on the basis of a sample of this size to lie below X_n. Repeating this for various assigned values of X_n, we may construct any part of the distribution of X'_K desired. Since n is small, the number of combinations of a and b to be investigated is limited. The labor involved in this procedure, while not small, will still usually not be prohibitive. This is especially true if, as will many times be the case, all that is desired is a specific test of the significance of the difference between the percentile and some other single value of X.

The setting up of a single form summarizing the above procedure is difficult and unwieldy in the general case for any percentile. This has been done in the special case of the medium, but the resulting expression is not given in the text because of its complexity and limited usefulness. The small sample case must be regarded as bordering on the trivial, because in extremely few instances are percentiles used in conjunction with samples of very small absolute size.

Samples of Intermediate Size. In some cases, a sample may be too large to permit convenient evaluation by the special procedures suggested in the text immediately above, but too small to permit the application of equation (17), especially in view of the simplifications incorporated in the latter which are based on the assumption of very large sample size. Under these conditions, the procedure suggested in the text following equation (13) will usually suffice. This method is most useful if all that is desired is a test of the significance of the difference between an estimated percentile and some preassigned quantity.

The validity of the use of this procedure depends upon how closely a linear function of one or more hypergeometric variates (depending on the type of sampling employed) may be represented by the normal probability distribution. In the case of the more central percentiles a fair agreement is obtained even with relatively small sample sizes.

Sample Allocations in Stratified Samples. Neyman and others have studied the question of the proper allocation of a limited number of schedules between strata to produce the most efficient estimate of the mean. It may be of interest to apply this same technique to the prob-

lem of estimating a percentile. First, it is assumed that the cost of obtaining each schedule is the same throughout the population. The total number of schedules to be obtained (total sample size) is represented by n_o. We then define a function V as follows:

$$V = [C^2K(1 - K)/F^2c]\sum_{i=1}^{r} N_i^2 p_{a_i}/n_i + L\left(\sum_{i=1}^{r} n_i - n_o\right) \qquad (20)$$

where L is an arbitrary Lagrange multiplier. The function V is differentiated with respect to n_i and n_j and the results equated to zero. Between the resulting two equations L is eliminated. From the result, it may be shown that to minimize the variance of our estimate of a percentile, the size of sample to be taken in any stratum should be allocated in accordance with the following relationship,

$$n_i = n_o\sqrt{N_iA_i} \Big/ \sum_{i=1}^{r} \sqrt{N_iA_i}. \qquad (21)$$

It is interesting to observe that the allocation of the sample is independent of the particular percentile which is to be estimated. Moreover, if the percentile is to be determined for the whole population, A_i becomes equal to N_i, and the allocations are simply made proportionate to the total number of individuals in each stratum.

Illustration. It was mentioned in the introduction that the standard error of the median may be substantially less than that of the mean in certain distributions, especially those of an economic character. A concrete illustration of this may be of interest. The accompanying tabulation presents a distribution of 1,200 families according to their annual incomes. This distribution is patterned after a sample obtained by the Study of Consumer Purchases in a restricted section of New York City. It thus represents what may be found in practice.

Under $250........	10	$ 2,500 to $2,999...	157
$ 250 to $ 499...	12	3,000 " 3,499...	88
500 " 749...	26	3,500 " 3,999...	52
750 " 999...	49	4,000 " 4,499...	33
1,000 " 1,249...	89	4,500 " 4,999...	20
1,250 " 1,499...	104	5,000 " 7,499...	54
1,500 " 1,749...	125	7,500 " 9,999...	15
1,750 " 1,999...	132	10,000 and over....	18
2,000 " 2,249 ..	126		
2,250 " 2,499...	90	All incomes.......	1,200

Applying equation (19), the standard error of the median of this distribution is estimated to be about \$34. The usual methods lead to an estimate of the standard error of the mean of \$73. The range of uncertainty of an estimate of the mean based on this sample would then be about double the range for the median.

The difference in reliability may be exhibited in more striking fashion in terms of the greater size of sample required to get an estimate of the mean having a standard error of only \$34. It is readily found that to provide such an estimate a sample of approximately 5,500 families would be required, or a sample more than $4\frac{1}{2}$ times as large as that required to estimate the median with the same standard error.

PRICES AND WAGES*

By Walter G. Keim
Bureau of Labor Statistics

In approaching the subject of prices and wages, a distinction should be made between two different relationships. First, there is the relationship between the earnings of a worker and the cost of goods which he purchases with his earnings. Then, there is the relationship between wages as a cost element in the production of goods and the price of the goods. The first of these relationships involves the income of the worker as measured, for example, by his average weekly earnings. The second involves unit labor costs, of which average hourly earnings are an important, but not a precise, indicator.

This paper is intended to describe the major trends in wages and prices and to indicate the various factors which have had a tendency to disturb the relationships between wages and prices since the beginning of the war in Europe. Lack of time, of course, precludes detailed discussion and analysis.

An analysis of the first relationship, that of changes in workers' income as measured by average weekly earnings and changes in the cost of living, indicates that since the outbreak of the war the economic status of the typical wage earner in manufacturing and mining has been materially improved. The cost of living between August 1939 and October 1941 rose 12 per cent while average weekly earnings advanced 34 per cent. This amounts to an increase of 20 to 22 per cent in the real weekly earnings of factory employees. Since the very outbreak of the war this relationship has persisted. Between September 1939 and September 1940, weekly earnings mounted 7.4 per cent while the cost of living in large cities actually declined .2 per cent. Similarly, between September 1940 and October 1941, the respective changes were an increase of 23.7 per cent in earnings and an advance of 9.7 per cent in living costs. It is likely that total annual earnings, would show an even greater increase because of fuller employment.

* A paper presented at the 103rd Annual Meeting of the American Statistical Association, New York, December 28, 1941. Since this paper was prepared, there has been some shift in the focus of the discussion regarding the relationship between prices and wages. The emphasis in this paper is upon wages as an element in the cost of producing goods and their consequent impact upon the price structure. Since the tremendous increase in the pace of war production which has occurred after Pearl Harbor, and the progressive reduction in the volume of goods available for the civilian market, there has been increasing concern with the relation of wages to total public purchasing power. During the summer of 1942, when this paper was going to press, primary emphasis in the formulation of national policy was upon this latter phase: upon the possibility that widespread wage increases, by increasing consumer demand, might lead to excess pressure upon the price structure.

To this observation that the status of the average wage earner has been materially improved, there must, however, be added a number of reservations: in interpreting these statistics it should be borne in mind that a significant proportion of the upward movement in average weekly earnings for all manufacturing industries is to be accounted for first, by a shift in employment. Under the impact of the defense program, *employment in durable goods industries*, which pay relatively higher wages, has risen more rapidly than in the nondurable goods field.

The figures just presented are *national averages* and workers have by no means been uniformly affected. In many lines of work the change in average weekly earnings has been much smaller than the average; for example, compared with the increase of 34 per cent in all manufacturing industries, weekly earnings have changed only 8 per cent in newspaper plants, 11 per cent in hosiery manufacturing. Moreover, the data for the country as a whole are not necessarily typical of defense areas where both the advances in earnings and the increase in cost of living have been greater than average.

Much of the increase in average weekly earnings represents longer hours of work resulting from the transition from part-time to full-time work and more recently the rapid lengthening of hours of work so that overtime at high rates is paid.

Despite these reservations, however, there can be little doubt that the average industrial worker is today much better off than he was before the war began.

We can turn now to the second relationship—that between wages as an element in costs of production and the prices of goods produced. There has been much talk about the role that wage increases have played in the advance in prices since August 1939. The figures available seem to indicate rather clearly that until the summer or autumn of 1941 at least, increases in wages have not been the primary factor in causing prices to go up.

A first indication that industrial wages have not been the key factor in price increases may be taken from the following figures contrasting the per cent of increase among different groups of prices. Whereas the all-commodities index of wholesale prices advanced approximately 23 per cent from August 1939 to November 1941, prices of basic raw materials, as shown by the Bureau's index of 28 basic commodities, increased 55 per cent over the same period; the index of all raw materials has gone up about 30 per cent; semi-manufactured goods have gone up about 20 per cent; finished manufactured goods about 10 per cent. The durable heavy goods, which are most affected by the defense program,

have gone up only about 15 per cent. The index for all commodities other than farm products, which advanced 19 per cent, is the one most comparable with the data on industrial earnings.

These figures show that the greatest price increases in this list since August 1939 have been among the raw materials. Wages do not constitute a key factor in the cost of any of these commodities. The greater the price advance, the smaller has been the element of fabrication by factory workers in this country. The converse is also true. The smallest price advance for any of the groups of commodities just cited is that for durable manufactured goods. It is in this field of durable goods that the largest wage advances have occurred. For durable goods as a whole, average hourly earnings rose about 28 per cent between August 1939 and October 1941, while the corresponding increase in the case of nondurable goods amounted to less than 12 per cent. Such increases in wages as have been granted obviously have been a minor factor in the price advances that have actually occurred. Price increases at wholesale have generally preceded increases in wage rates and have far outrun them.

Wages, of course, constitute only one of the costs of manufacturing and hence the selling prices of manufactured goods do not have to rise at the same rate as labor costs, merely to cover the cost of any given wage increase. When all manufacturing industries are considered together, account must be taken of the fact that the output of one producer frequently becomes the raw material of another. Hence, the question is: what is the ratio of wages paid to the value added by manufacture? Census data show that wages constitute about 40 per cent of value added by all manufacturing operations. On the average, therefore, a 2½ per cent increase in selling prices will cover the cost of a 6 per cent general wage increase provided that output per man hour is unchanged.

Let us analyze the actual sequence of events. There have been two major upswings in prices since the beginning of the war in Europe. One occurred in the latter part of 1939 when the index for all nonagricultural commodities increased 6 per cent from the outbreak of the war at the end of August to a peak in December of that year. During this period, average hourly earnings changed hardly at all. There was a general reduction in prices in the early months of 1940, but after August 1940, when the effects of the American defense program began to be felt, prices resumed their upward trend, and have been rising almost without interruption ever since. From August 1940 to March 1941, prices of all commodities rose on an average by 8 per cent and average

hourly earnings rose by 4 per cent. In each of these two periods of rising prices and wages from the outbreak of war to March 1941, therefore, the advance in prices cannot be attributed to higher wages.

It was not until March 1941 that important changes in the wage rates occurred. The major wage advances in early 1941 really began in 4 major industries—cotton manufacturing, coal, steel and automobiles. In each of these industries it was evident that profits had increased. Since that time, there has been a general fanning out of these wage increases to other industries.

In the case of cotton goods, wholesale prices had risen by 13 per cent in the 5 months prior to the wage increase (October 1940 to March 1941) and actually by 4½ per cent during March alone. During this period, moreover, prices of raw cotton were advancing only moderately and unit overhead was decreasing because of rapidly advancing sales. Mill margins on cotton cloth during this period had risen by an average of 30 per cent to the highest levels on record.

The wage increase in cotton manufacturing that began in March amounted to 10 per cent in the North and 7 to 8 per cent in the South among those cotton mills that made the adjustment. The average effect from March to June was to raise hourly earnings in cotton mills as a whole by 6.6 per cent. In July 1941, hourly earnings advanced again as a result of adjustments to the new 37½ cent legal minimum wage of the Fair Labor Standards Act, bringing the total increase from August 1939 to September 1941 to 26 per cent. Assuming for the moment that labor costs rose by the same percentage, the added cost could have been met by a 6 or 7 per cent increase in selling prices, since direct wages represent only about one-quarter of the value of cotton-mill products.

Not only had prices already risen by 13 per cent between October 1940 and March 1941, but they continued to advance after March. By November 1941, the average price of all cotton goods was 61 per cent higher than in August 1939 and 47 per cent above the level of October 1940. Even though raw cotton prices advanced sharply since February 1941, the mill margins in October 1941 were 79 per cent higher than in August 1939. On this advance, as we have seen, only a very small proportion could be attributed to a rise in wage rates.

The wage increase in steel mills and in some of the leading automobile factories amounted to 10 cents an hour. This would be equivalent to an 11½ per cent rise in steel wages and about a 10 per cent rise in the automobile industry if univerally applied. By July, hourly earnings in steel mills had leveled off at 11 per cent above the average for March. A number of the new wage agreements in the automobile industry pro-

vided for increases of 5 cents or 8 cents in place of the 10 cent advance which was granted by General Motors on May 15, and by July the average increase for the industry over the level of March amounted to 8 cents an hour. Since the wage increase in steel mills, the quoted price of steel has shown no significant change. Automobile prices were raised in October, reflecting in part changes in unit cost resulting from a combination of factors, shifts to defense production, changes in raw material costs and labor. Looking at the overall picture during the period March 1941 to October the price index for nonagricultural commodities went up 14 per cent and average hourly earnings increased 10 per cent.

In interpreting all these figures, it cannot be overemphasized that higher average hourly earnings do not necessarily reflect wage rate increases. Thus, if a man is being paid on a piece basis, he may increase his hourly earnings simply by producing more per hour. Similarly, higher hourly earnings may reflect increased overtime at premium rates rather than a change in basic wage rates. More broadly, higher hourly earnings do not necessarily mean rising unit labor costs. The cost of labor per unit of output—that is, per yard of goods or per ton of coal or per ton of copper—depends on how much is produced per hour as well as on what the worker gets per hour. While hourly earnings have been advancing in recent years, so has output per man-hour. Indeed, the data at the disposal of the Bureau of Labor Statistics indicate that the output of manufacturing industry per man-hour increased about 16 per cent on the average from 1937 through the summer of 1941. In the same period average hourly earnings increased by 17 per cent; hence, unit labor costs advanced only 1 per cent despite the wage increases since March. In other words, the increase in hourly earnings was offset almost entirely by the increase in the amount of goods produced per hour.

So much for the immediate past. We have seen that higher wages have not been primarily responsible for the rising general level of prices. But at this time many factors are tending to disturb wage-price relationships. Up to the present, in many industries, increases in actual unit labor costs have been more than offset by lower overhead per unit. However, from this point on, there will be definite limits to our ability to increase productivity.

Similarly, we are reaching the limit of our capacity to expand production. There will be numerous shifts from consumer goods production to durable defense goods production. The period of increasingly lower burden per unit of production is reaching an end. In fact, some industries will be forced to curtail production and, depending on the nature

of their output, they may not be able to transform their plants satisfactorily in order to produce defense goods. Hence, many will be faced with the problem of vast curtailment of normal production with the attendant increases in unit overhead costs.

These shutdowns may be due to a number of causes, of which the most important will probably be shortages of materials. The automobile industry, the electric refrigerator industry, the washing machine industry, in fact nearly all industries producing consumers' durable goods, are going to be seriously affected in this way. Also, many manufacturers of perishable goods will find that they will have difficulties in replacing manufacturing equipment, because of the shortages of metals used in making the equipment.

Certain producers may well be faced with such problems as power shortages. Thus, during this last year, production costs in the Southern textile industry increased because of the inability to secure sufficient power. Then there will be the question of transportation difficulties, shortages in certain types of skilled labor, and similar problems.

Thus in certain industries operating at capacity or under limitations imposed by the war program further wage increases will inevitably bring a demand for higher prices.

This means in the long run the raising of living costs. Even at present, the large increases in prices in wholesale markets, particularly for foods (not due to wage increases) are rapidly filtering through to retail markets and are raising the cost of living. There will be many new demands for increases in wage rates. In the last war unions occasionally included in their contracts an escalator clause providing for increases in wages with advances in living costs or prices of the product they produced. At the present time, however, unions are generally not requesting this provision. Although there are exceptions more often they merely use the record of the upward changes in cost of living to justify their demands for increased wages.

Unless careful consideration is given to the ability of concerns to pay higher wages without further increases in their selling prices, we may well face a cycle of wage and price increases that will become a serious threat both to the general welfare of the Nation and to the conduct of the defense program itself.

THE USE OF TESTS OF SIGNIFICANCE IN AN AGRICULTURAL EXPERIMENT STATION*

By George W. Snedecor
Iowa State College

The purpose of this paper is to recount some of the uses made of tests of significance in our research procedure. In order to clarify my presentation it is necessary at the very outset to specify rather carefully the circumstances in which these tests are applied.

Our samplings are of two types, the rather intensive sort involved in controlled experiments and the more extensive surveys conducted by questionnaire, visitation, or the examination of extant records. The objectives are the same; namely, unbiased estimate of population facts and probability statements based on these estimates.

The populations sampled are ordinarily described in characteristic fashions. Those subjected to experimentation are specified under some such caption as "Materials and Methods." They are often voluntarily restricted by the choice of experimental animals and plants as well as by the attempted control of extraneous effects. For example, the rat may be selected for testing the difference in weight gain consequent upon two treatments. The animals used may be taken from some highly inbred colony maintained in a carefully controlled environment, and the sampling confined to individuals of the same age and sex. Thus, the heredity, the environment and the two treatments of the chosen animal all enter into the specification of the population. The population sampled by inquiry must be described no less meticulously. The region sampled, the questions asked, the sampling unit together with the mode of its selection, the stratification adopted, and the date are all pertinent.

Coming now to the immediate topic of my paper, I remark that an experiment is a sampling designed wholly or in part to test some null hypothesis. To fix attention, let us think of a field plot trial to discover if the application of a specified amount of fertilizer affects the yield of a certain crop. The objective is attained if we learn whether or not the use of the fertilizer differentiates the sampled aggregate into two populations; the one, yields from treated plots, and the other, yields from untreated. You see at once that either of two null hypotheses may be suitable. If the fertilizer applied is supposed to cause a

* A paper presented at the 103rd Annual Meeting of the American Statistical Association in joint session with the Institute of Mathematical Statistics, New York, December 27, 1941.

more or less durable improvement in the soil, its cost may be charged to future gains, and the appropriate null hypothesis is the identity of the yields on fertilized and unfertilized plots. If, on the contrary, the value of the treatment lies only in its effect on the current crop, the hypothesis may be that the difference in yields is exactly sufficient to pay the cost of the application.

With the experimental data in hand the investigator calculates the mean difference in yield, together with some appropriate quantity such as t, z, or F. The probability of this quantity is the so-called test of significance. It is really evidence concerning the chosen null hypothesis, and must be considered as part of all the other evidence accumulated prior to and during the experiment. In the light of his entire experience, including the test of the pertinent hypothesis, the agronomist then makes a decision about the efficacy of the fertilizer under the experimental conditions obtaining.

Of great utility in our practice is an extension of the experiment just described to include comparisons among large numbers of varieties or treatments. The data from such an investigation lead to a set of means whose variance is comparable with a second variance provided by the sampling, commonly designated as experimental error. To test the null hypothesis, the ratio of these two variances is calculated and compared with the tabulated distribution. A large value of the variance ratio, F, is looked upon by the experimenter as evidence that his mean yields are more widely dispersed than would be expected in random sampling from a single population. If this evidence is compatible with that already available he concludes that the higher yielding varieties, other characteristics being satisfactory, are worth selecting for further trial, or for recommendation to farmers operating in an environment similar to that of the experiment. On the other hand, a small value of F warns him that the population of yields represented by these varieties may be undifferentiated. If this is in conformity with other experience he may ignore yield as a criterion, freely selecting such varieties as may possess some other desirable characteristics.

As a third example of our use of tests of significance, I have chosen another often encountered feature of yield tests; namely, the effect of stand; that is, the number of plants per unit of area. This is a variate concomitant to yield, difficult to control but easy to measure. It may be a variety characteristic, or it may be due to random variation in germination and early growth, the distinction being important agronomically. If the variety variation in stand is found to be significant, an otherwise promising kind may be rejected because of poor seed germi-

nation. On the contrary, a small F may lead the plant breeder to ignore stand as a basis for selection, and to evaluate yield by use of expected mean values adjusted by the regression of yield on stand. In this latter event, he may increase the precision of his evaluation by eliminating from experimental error the measured variance attributable to the random irregularities of stand.

Quite a different use of tests of hypotheses is that commonly invoked if experimental error rests on the variation of individuals within two or more groups receiving different treatments. In order to test the hypothesis that the means are drawn from a common source, the group variances must be pooled. If these batches of variance are samples from populations with different variances the F-test doesn't give unambiguous information about the means. Hence, the hypothesis of homogeneity of variance is set up and tested. An unusually large value of chi-square may lead the investigator to use some appropriate transformation of his data in order that the F-test may give him desired evidence about the variation of the group means.

The foregoing tests have been selected partly because of their common occurrence and partly because they illustrate some of the types of decisions based on experimental evidence. In the extensive type of sampling, tests of significance may not be used at all. If an estimate of population total is the statistic desired, the appropriate statement about probability is the fiducial limits. The effectiveness of the sampling design is ordinarily examined by studying the efficiency of stratification and of the size and allocation of the sampling unit. Nevertheless, group means of such extensive samples are often compared, and regression together with its linearity must frequently be tested. Hence, it is not far from the truth to say that a majority of the steps in the examination of our data are guided by tests of some null hypothesis.

I should like to make it clear that these tests may not be made formally. Careful perusal of the data occasionally obviates the necessity of calculations. Again, since careless design or faulty execution of an experiment may yield no estimate of error, the statistician may be forced to resort to rough approximations based on knowledge of range or coefficients of variation. Naturally, he fortifies his decisions about significance with greater than the usual factor of safety.

You may ask if conclusions and recommendations flowing from an experiment are ever based wholly on the probability of the testing statistic. I suspect not, since no experiment can yield proof but only evidence, and since few investigators design a sampling without an extensive background of information and experience. Moreover, there

is nearly always evidence gained during the progress of an experiment which is not incorporated into the numerical data. Conceivably the investigator may come to a final trial with pros and cons exactly balanced, leaving his decision to the outcome of the experiment. If so, what odds would he demand against the null hypothesis to reject it? Greater than 99:1, I suppose. More often an experiment is conducted to provide a link in some long chain of evidence. I see no necessary objection in such case to rejecting the hypothesis on the basis of any odds against it greater than 1:1. Except for beginners and those who are easily swayed by their emotions, I think there is little merit in setting arbitrary values to probability, then decreeing that smaller must lead to rejection. The 5 per cent and 1 per cent points are convenient milestones which the investigator will note in passing, but any probability turned up constitutes evidence pertinent to decisions.

All the circumstances of applied sampling warn us against blind subservience to any conventional probability. The distribution of a testing function is worked out from assumptions of normality of population, randomness of sampling, independence, etc. This model is probably not exactly reflected in any actual sampling. It is known that rather generous relaxations of some of the conditions have little effect on the probabilities involved. On the other hand, work of a number of investigators makes it clear that inaccuracies of several per cent either way from the 5 per cent point may not be uncommon. This does not disturb the sampler who looks upon the test of his hypothesis as contributory evidence, though it may be devastating to the worshipper of 5 per cent. It is the consulting statistician who finds arbitrary points handy because his decisions must be based rather solidly upon the evidence of the data themselves.

I am coming to believe that the term, "test of significance" is creating more confusion than it resolves. The phrase is not descriptive of the logical and experimental concepts involved. Partly for this reason, the dictionary definitions of the words blind many to their meaning in terms of probability. Perhaps the time has come to state the probability of the testing function as the end-point of the statistical investigation leaving the researcher to combine this evidence with that already accumulated, then to rest his decision on the whole of it.

MECHANIZATION OF STATISTICAL DRAFTING

By R. von Huhn

During World War I statistical drafting came into its own. The various war agencies were in constant search for men who were familiar with the technique and principles of graphic presentation. Needless to say, most of the men who were qualified came from the engineering profession. Brinton, the engineer, had published his book on graphic presentation in 1914, the first of its kind.

In the years following, the late Knoeppel became one of the foremost exponents of graphic production control.[1] He devised large control boards which, by mechanical means, served to guide the production engineer as to the status of balanced inventory, schedules, actual output and deliveries. Use was made of the principle of the "Gant Chart," translated into a graphic mechanical set-up by means of movable strips, replaceable pins, etc. The devices were kept in the office of the supervising production engineer for purposes of immediate control. The same method is used by various industries in the present emergency to control the flow of production and assembly of implements of war.

In 1917 there were at the disposal of the graphic analyst only few materials which enabled him to perform his work efficiently. Since then the situation has greatly changed. New reproduction methods have been discovered and others have been perfected. New materials such as Zipp-a-tone, acetate overlays, transparent tape, rubber cement, scotch tape, floating letter type, photo-letter type, stack type letters, the Wrico and the Leroy mechanical lettering sets are now on the market.

Although only indirectly related to the subject matter discussed here, mechanization of computing operations should also be mentioned. The hand operated computing machine, while still in use, has given way to the electrically operated machine. The various models on the market have been constantly perfected, keeping pace with the requirements of the statistician or accountant. The automatic tabulating machines which are essential for the mass production of statistics need hardly be mentioned here. Incredible as it seems, each year new improvements are made and the older designs are revised.

At present, as far as the writer knows, there is only one mechanical device on the market which eliminates drafting entirely—the so-called

[1] C. E. Knoeppel, *Graphic Production Control*, pp. 477, Ill. New York, McGraw-Hill Book Company.

Cosmograph—manufactured by the International Business Machines Corporation. The device is used to show the flow of a percentage distribution.[2] The other device designed by the writer for the use of the United States Government is the mechanical intensity shading map which was described in the December 1938 issue of this JOURNAL. The writer has recently constructed a crude demonstration model of a mechanical bar chart which takes care of percentage bar diagrams showing four different component parts, and their respective shadings adjustable to any position.

The advance in technique has, on the whole, been one-sided with emphasis on the perfection of materials for reproduction methods. This may be termed partial mechanization of statistical drafting. Complete mechanization of drafting may be defined as a process by which a final chart or map is produced entirely without the aid of drafting instruments in the same manner as the typist prepares a copy. It is interesting to note that H. Gray Funkhouser,[3] who has made what is, up to the present, the most thorough contribution to the history of graphic presentation, did not discover in his researches evidences of mechanization of statistical drafting, nor did he mention its possibilities.

Complete mechanization of statistical charts and maps will never answer all the demands set forth by the investigator who makes use of graphic presentation; nor can it take the place of individual originality which plays a great part in the design of graphs analyzing special situations. The United States Civil Service Commission recognizes now that graphic presentation and analysis are well-defined professional fields requiring special qualifications and training.

There is, however, a distinct place for a machine which will produce simple line charts especially time series which show long term trends. The job of plotting a 20-year trend by months for three or more curves is a tedious one and belongs, in the writer's opinion, to the field of mechanization.

What then are the requirements for a statistical charting machine? The following are some of the points which seem essential:

1. Standard 10-keyboard from 0 to 9, similar to the one found on the Sundstrand Adding Machine.

[2] The above statement does not refer to mechanical devices used in industry which automatically keep graphic records of temperature, gas consumption, variation of speed, barometric pressure etc. The roles which these instruments play in our industry have been described in a monograph entitled: "Industrial Instruments and Changing Technology," by G. Perazich, H. Schimmel and B. Rosenberg, W.P.A. National Research Project, Report No. M-1, October 1938, Philadelphia, Pa.

[3] H. Gray Funkhouser, "Historical Development of the Graphical Representation of Statistical Data," *Osiris*, Volume III, Part I, 1937, Bruges (Belgium). Funkhouser's fundamental work should be read by all students interested in graphic methods of presenting statistical data; besides an interesting text, it contains an excellent bibliography.

2. Adjustable vertical and horizontal movement—the latter designed to permit variation in distance between ordinates, as the time scale may require.

3. The most important part of the machine is the printing mechanism, which controlled by the keyboard would coordinate the desired horizontal and vertical movement, thereby executing automatically what is otherwise left to the draftsman. This mechanism or *integrator* is visualized to run on tracks by tooth and pinion and to respond vertically to the operation of the keyboard in such manner that the machine-plotted points are either arithmetically or logarithmically proportional in height in relation to a common reference line. The integrator should have a flexibility which permits scale variations according to the range of data which are to be plotted. The integrator then is to have two principal functions, namely, first to locate a point to be plotted in relation to the required scale and, secondly, to connect the point by a line with the next point plotted, in order to produce the required trend line. The printing mechanism should also be designed in a manner that when two or more trends are plotted, each will show a distinctly different legend or pattern, i.e., solid line, dash line, etc.

If the mechanical requirements of the integrator, as suggested above, involve too many mechanical obstacles, a compromise would be to design the integrator to take care of plotting only by locating the points correctly, which then could be connected by the draftsman. While this would not be complete mechanization, it would be immensely helpful in speeding up the completion of a long-term trend chart. The plotting points printed by the integrator would have to differ slightly, as for example—small solid circles, small open circle points, etc., thereby indicating to the draftsman that they belong to the same series.[4]

Due to the manpower war requirements the scarcity of competent men trained in the field of statistical drafting becomes greater every day. It seems that such a device as the writer visualizes should have been on the market long ago and that its construction would have occurred to the great business machine companies. One reason for its failure to appear may be the fact that the market for a machine of this nature would be fairly limited.

The question now arises: What would be the advantages if a charting

[4] The solution of the mechanical *integrator* will tax the imagination of the designing engineer. The vertical movement may be activated through the use of calibrated cams, controlled by the keyboard operation or by cylindrical "wedges" or inclined planes attached to drums, similar to those which control and activate the motions of automatic machinery, or the solution may be found in an electrically controlled mechanism such as we find in tabulating machines. Still another direction of solution may be found in the application of the photoelectric cell which plays an increasingly important part in the new design and perfection of instruments for the automatic precision control of industrial processes.

machine, such as proposed in the preceding paragraphs, were available? Doubtless some of the following can be numbered:

1. It would be possible to turn over routine line charts to an operator untrained in drafting.

2. The tedious plotting of points would be eliminated.

3. The time for producing line charts would be decreased to a fraction of that previously required and consequently a greater volume could be turned out per man-hour.

4. The graphs would have a uniform appearance and could be subjected to photostat or multilith reproduction.

5. The cost of charting would be decreased and the machine would pay for itself in a relatively short time.

6. There is strong possibility engineers would want so to design the machine that it could be connected electrically with the tabulating machine unit, with the result that the keys of the mechanism would be operated automatically and simultaneously through the punched machine cards. Thus the process of counting and graphic presentation of time series could be accomplished during the same "run" of cards.

The above are some of the advantages which would be derived if such a machine were available. In conclusion, it should be stated that mechanization of statistical drafting is still in its infancy but sooner or later progress will be made in that direction.

COMMITTEE ON NOMINATIONS

President Lotka has appointed the Committee on Nominations for 1942. The Committee consists of Frederick F. Stephan, Cornell University, Ithaca, New York, Chairman; Joseph S. Davis, Food Research Institute, Stanford University, California; and George W. Snedecor, Iowa State College, Ames, Iowa. The report of the Committee on Nominations will be published in the November Bulletin.

R. L. Funkhouser, *Secretary*

SKEWNESS OF COMBINED DISTRIBUTIONS

By Theodore E. Raiford
University of Michigan

It often happens that investigators, although eager to publish the results of their investigations, jealously guard the basic data from which the results were obtained. Since accumulative experience tends towards conclusiveness, combined results from several investigators on the same problem would often be of very great value. While making a study of the combined results from several investigations two points have arisen which seem worthy of special presentation, namely, (1) a new formula for determining the measure of skewness for the combined (parent) sets of data in terms of parameters which characterize the individual sets (subsets) of measurements, and (2) that combining subsets all of which are skew in the same direction does not necessarily result in a parent distribution with skewness in that same direction. This short paper presents a development of the formula suggested in (1) and a numerical example which not only illustrates its use but also establishes the truth of statement (2).

Among the things usually given in the statistical part of a report on one's findings from a single set of data, in addition to the number of measurements, are the mean, the standard deviation, and a measure of skewness. We shall denote these parameters for the parent distribution by the usual letters, without subscripts, namely, N, M, σ, and α_3, and the corresponding parameters from the i-th subset we will designate as n_i, M_i, σ_i, and $\alpha_{3:i}$. Moreover, as the measure of skewness we shall use the third moment about the mean, expressed in standard units.

A formula for the mean of the combined distribution from k subsets has been presented in many places in the form

$$M = \frac{\sum_{i=1}^{k} n_i M_i}{N} .$$

In this notation, $N = \sum_{i=1}^{k} n_i$. One of the most convenient forms for the computation of σ is one making direct use of the parameters contributed from the subsets and based on the formula,[1]

$$\sigma^2 = \frac{\sum_{i=1}^{k} n_i(\sigma_i^2 + d_i)}{N}, \quad \text{where} \quad d_i = M_i - M.$$

[1] H. C. Carver, *Notes on Elements of Mathematical Statistics*, Edwards Brothers, Inc., 1930.

Forms for the computation of skewness of combined distributions have not been so numerous,[2] but the notation of the above formula for σ suggests also a concise form in which α_3 may be expressed.

In a distribution in which X denotes the variable, using the notation of serial distributions for simplicity, we have the following relations,

$$M = \frac{\sum X}{n}; \qquad \sigma^2 = \frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2; \quad \text{and}$$

$$\alpha_3 = \frac{n^2 \sum X^3 - 3n \sum X^2 \sum X + 2(\sum X)^3}{n^3 \sigma^3} \tag{1}$$

where the summations are from 1 to n. From these relations we get

$$\begin{aligned} \sum X &= n\, M, \\ \sum X^2 &= n(\sigma^2 + M^2), \qquad \text{and} \\ \sum X^3 &= n \cdot \sigma^3\, \alpha_3 + 3n\, \sigma^2\, M + n \cdot M^3. \end{aligned} \tag{2}$$

If now we denote the measurements in subset No. 1 by X_1, those in subset No. 2 by X_2, and so on, we have at once from definition,

$$\begin{aligned} N \cdot \sigma^3 \cdot \alpha_3 &= \sum (X_1 - M)^3 + \sum (X_2 - M)^3 + \cdots + \sum (X_k - M)^3 \\ &= \sum X_1^3 - 3 \sum X_1^2\, M + 3 \sum X_1\, M^2 - n_1\, M^3 \\ &\quad + \sum X_2^3 - 3 \sum X_2^2 \cdot M + 3 \sum X_2 \cdot M^2 - n_2 \cdot M^3 \\ &\quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\quad + \sum X_k^3 - 3 \sum X_k^2 \cdot M + 3 \sum X_k\, M^2 - n_k \cdot M^3 \end{aligned} \tag{3}$$

where the summations in X_i are from 1 to n_i. If the summations in this equation are replaced by equivalent expressions defined by the relations given in (2) and if like terms are collected, the above equation becomes

$$\begin{aligned} N \cdot \sigma^3 \cdot \alpha_3 = {} & n_1[\sigma_1^3 \cdot \alpha_{3 \cdot 1} + 3\, d_1 \sigma_1^2 + d_1^3] \\ & + n_2[\sigma_2^3 \cdot \alpha_{3 \cdot 2} + 3\, d_2 \sigma_2^2 + d_2^3] \\ & \cdots\cdots\cdots\cdots \\ & + n_k[\sigma_k^3 \cdot \alpha_{3 \cdot k} + 3\, d_k \sigma_k^2 + d_k^3]. \end{aligned} \tag{4}$$

From this equation a formula for α_3 may therefore be written as[3]

$$\alpha_3 = \frac{\sum n_i \sigma_i^3 \alpha_{3 \cdot i} + \sum n_i d_i (3\sigma_i^2 + d_i^2)}{N \sigma^3}. \tag{5}$$

[2] W. D. Baten, "A Formula for Finding the Skewness of the Combination of Two or More Samples," this JOURNAL, March 1935, pp. 95–98.

[3] T. E. Raiford, "The Measurement of Variations in Asymmetrical Data with Applications to Hemoglobin Statistics of Infants," *Human Biology*, February 1938, p. 130.

A tabular form for the application of this formula is shown in the table in combining the results from 13 subsets of data presented from studies over a period of 13 months of the amount of hemoglobin per 100 cc. of blood in a certain group of children. Other columns may be added for convenience, depending mainly upon what mechanical device is used to perform the computations.

TABLE I

Subset	n_i	M_i	σ_i	$\alpha_{3:i}$	d_i	$\sigma_i^2+d_i^2$	$n_i\sigma_i^3\alpha_{3:i}$	n_id_i	$3\sigma_i^2+d_i^2$
1	108	13.33	1.92	−.31	1.98	7.0008	−308.010	332.04	14.0700
2	272	11.17	1.31	−.35	−.18	1.7485	−214.018	−48.00	5.1807
3	329	10.92	1.01	+.10	−.43	1.2008	70.315	−141.47	3.4207
4	332	11.34	1.00	−.04	−.01	1.0001	−13.280	−3.32	3.0001
5	308	11.60	1.08	−.30	.25	1.2280	−130.007	77.00	3.5017
6	312	11.67	1.05	−.54	.32	1.2040	−105.030	99.84	3.4000
7	201	11.45	1.00	−.52	.10	1.0100	−151.320	20.10	3.0100
8	273	11.30	1.03	−.40	−.05	1.0634	−110.320	−13.65	3.1852
9	210	11.20	1.10	−.37	−.15	1.4330	−153.383	−30.00	4.2708
10	203	11.16	1.07	−.44	−.10	1.1810	−141.762	−40.07	3.4708
11	253	11.03	1.17	−.26	−.32	1.4713	−101.302	−80.00	4.2001
12	210	10.98	1.11	−.55	−.37	1.3090	−157.001	−77.70	3.8332
13	198	10.98	1.14	−.59	−.37	1.4308	−173.073	−73.26	4.0357

$\Sigma n_i(\sigma_i^2+d_i^2)=5457.1550$; $\Sigma n_i\sigma_i^3\alpha_{3:i}=-1858.430$; $\Sigma n_id_i(3\sigma_i^2+d_i^2)=3027.5243$
$N=3455$; $M=11.35$; $\sigma=1.20$; $\alpha_3=+.20$

The example was chosen as an application of the formula given in equation (5) as it is almost a perfect illustration of the second point made. As a matter of fact, if subset 3 (the only one with a positive skewness) is omitted, the measure of skewness for the parent distribution formed by the remaining 12 subsets is found to be $\alpha_3=+.20$, thus establishing the truth of statement (2), the second object of this paper. It is not the purpose here to discuss the conditions under which subsets may properly be combined, but the example does call attention to the kind *of erroneous statement* which *may easily* be made in a general statement.

Since α_3 is positive or negative according as the numerator of the formula given by equation (5) is positive or negative, let the numerator be written in the form $\sum n_i\sigma_i^3\alpha_{3:i}+\sum n_i{}^-d_i(3\sigma_i^2+d_i^2)+\sum n_i{}^+d_i(3\sigma_i^2+d_i^2)$, where ${}^+d_i$ corresponds to the difference in which $M_i>M$ and ${}^-d_i$ corresponds to the difference in which $M_i<M$. The quantity $(3\sigma_i^2+d_i^2)$ is obviously always positive. If then the $\alpha_{3:i}$ are all of the same sign, say negative, α_3 for the combined distribution will be positive or negative according as $\sum n_i{}^+d_i(3\sigma_i^2+d_i^2)$ is greater than or less than $\sum n_i\sigma_i^3\alpha_{3:i}$ $+\sum n_i{}^-d_i(3\sigma_i^2+d_i^2)$. A similar statement is of course true in case the $\alpha_{3:i}$ are all positive.

BOOK REVIEWS

Glenn E. McLaughlin
Review Editor

Three Aspects of Labor Dynamics, by W. S. Woytinsky. Washington: Committee on Social Security, Social Science Research Council. 1942. xiv, 240 pp. $2.50.

This is a statistical investigation into certain of the changes in the composition of the labor force—principally labor turnover in prosperity and depression, the duration of individual employment, and net accessions to the labor market under the influence of business depression. It is largely an expansion of three earlier summaries published for the Social Science Research Council and the Social Security Board.

The material on labor turnover, constituting Part I of the volume, reviews the available statistics, footnotes the special studies, and details the history of this field of labor statistics since it was articulated during the period of the first world war. Turnover rates are analyzed in relation to business fluctuations, seasonal variations, types of industries, localities, skill, sex, age, working conditions, and length of service. An attempt is made to isolate the "unstable" element in the labor supply. The well-documented conclusions concerning economic behavior under these varying conditions and circumstances are fairly obvious.

Part II assembles available material on the turnover of the unemployed according to duration of inactivity in an effort to determine the "hard core" of unemployment. This area of labor statistics has been studied more carefully in recent depressions because of its pertinence to problems of relief, reemployment measures and social security. The author is, however, not concerned with measures of relieving unemployment. His discussion is limited to an explanation of the estimates themselves.

Part III deals with the phenomenon of supplementary earners in the labor market during depressions and contains interesting material on family unemployment, including an analysis of the Philadelphia unemployment studies. The author presents evidence and interprets available data to prove that the apparent larger unemployment in multi-worker families is due to the increase in the number of persons seeking employment when the chief earner is unemployed. The data permit such an interpretation, but this does not appear to be the only explanation. "Humane" employment policy of management has endeavored, where it is not too costly, to retain workers who are the sole earners in the family. Among the alternatives available, this discriminatory lay-off system has been considered good public policy. The principle was formally adopted by the Government in the automobile settlement of 1934.

It is a common phenomenon of the labor market that employment in-

creases with improving business; at such times additional workers in families may also be induced into the labor market. This is not to deny that in communities of diversified industries the differential demand for labor may bring the results which Mr. Woytinsky anticipates. The essential fact is, however, that there is a labor reserve of persons not actively in the labor market and that these persons—boys, young women, married women, and older "retired" persons—become searchers for jobs only under exceptional conditions, such as ease of getting jobs and the inducement of high pay in prosperity or the desperate need for cash in depression.

This sober treatment of *Three Aspects of Labor Dynamics* is based largely on studies in depressions. The "labor dynamics" of the past two years, involving a radical distortion of customary relations between workers and their jobs, are not considered. But, if "normalcy" returns again, the volume should be valuable as a summation of information for statisticians and economists on the perennial phenomenon of the relation of available jobs to the people who live by them.

GUSTAV PECK

Washington, D. C.

British Unemployment Programs, 1920–1938, by Eveline M. Burns. Washington: Committee On Social Security, Social Science Research Council. 1941. xx, 385 pp. $2.75.

This authoritative study of Britain's experience through the inter-war years of dealing with unemployment unquestionably will become a standard work for students of the subject in all countries. It is a good climax to a long series of excellent works on British experience with one of the knottiest and most universal problems of our times. Dr. Burns presents a careful and technical analysis of British experiments and achievements in the field. Furthermore, she evaluates the principles developed and basic problems involved in British experience with comprehensive knowledge and penetrating insight. The text is well documented and is accompanied by an admirable statistical appendix.

The study presents a careful analysis of the three periods since 1920 of the British system for alleviating the distress of unemployment; the attempt to expand unemployment insurance to provide benefits for both short and long term unemployment lasting from 1920 to 1931; the attempt from 1931 to 1934 to curtail unemployment insurance in the interest of financial solvency with "transitional" benefits for the long term unemployed which yet were distinguished from local public assistance; the program in effect since 1935 by which a highly integrated but dual system of unemployment insurance and unemployment assistance is established with financial responsibility for the latter provided from central tax funds. The author concludes that the last scheme is a notable achievement in providing maintenance for the unemployed which both reaches a rather high degree of coverage

and coincides relatively well with prevailing political principles as to the relation between government and the individual citizen.

The strengths of the present system (the word refers to times prior to the present war, obviously) lie in its relative stability, its fair success in meeting a basic and national minimum standard of life, and its administrative procedures. The achievements of the last are due in considerable part to the excellent program of local advisory councils and laymen's participation as well as to Britain's well-known civil service. The weaknesses of the system are found in part in the anomalous distinction between the short term unemployment insurance and long term unemployment assistance. Beyond any such distinction in principle, however, the author finds the most important flaw in British programs for dealing with unemployment is the failure to provide to any considerable extent for prevention of unemployment or adequate rehabilitation of the long term unemployed.

Dr. Burns concludes that the British have progressed far in the maintenance of the unemployed. They have done it by developing categorization of the unemployed (useful in Britain) and centralization of financial responsibility. The system has created a new relation between government and citizens with the responsibility of the former clearly accepted though the obligations of the latter and the social and economic functions of the modern family are still to be defined. The policy of prevention, the counselling and retraining of unemployed workers, the creation of work itself, whether by public works or by some measure of governmental control over industry and central planning has yet to be worked out in Britain as in other western countries.

The reviewer would have liked rather more attention given to the functions of the employment service in relation to dealing with unemployment. The reviewer also questions even more than the author seems to do the acceptance by British organized labor groups of the present system of maintenance. These, however, are minor criticisms. Dr. Burns' conclusion should be studied by all students of social, economic and political problems of modern industrial society.

MILDRED FAIRCHILD

Bryn Mawr College

Dimensions of Society, A Quantitative Systematics for the Social Sciences, by Stuart Carter Dodd. New York: The Macmillan Company. 1942. ix, 955 pp. $12.00.

The sub-title of this book, *A Quantitative Systematics for the Social Sciences*, accurately describes its content. It consists of a new and ingenious system of classification under rubrics transformable into numerical expressions. There are six parts: part one deals with the sectors of society, indicia (of traits), population, space, time; and the following parts elaborate in detail each of these sectors. The author has done a prodigious amount of work in assembling and classifying an enormous and highly varied mass of material. His

system is original and closely knit and, like all his work, shows rare capacity for painstaking analysis and organization.

The S-theory of Dodd's classification consists of a scheme of hypotheses which assert that every tabulation, graph, map, formula, prose paragraph (occasionally supplemented by the photograph of an individual, *vide* Gandhi, p. 86, or a composite of movie faces, *vide* p. 667) or other set of quantitative data in any of the social sciences may be grouped under its rubrics. The general sectors are four: T = time, L = space, P = population, and I = indicators of some trait. Thus $S = {}_s{}^s(T; I; L; P)_s{}^s$.

By using pre-and-post-superscripts and subscripts (for each sector), the author accounts for exponents (self-multiplication), sub-classes (aggregation), class-intervals, and *cases* (sometimes an individual item). Hundreds of illustrative specimens are considered and classified by this scheme. The heart of the S-theory (p. 41) is the quantic formula and the quantic number, since these provide (so the author claims) a thoroughgoing basis of classification for all quantifiable societal phenomena. Thus the quantic formula for the frequency distribution is, $S = T^0; I^1; L^0; P^1$, and the quantic number lifted off is, 0; 1; 0; 1. The quantic formula for an institution is, Ins. ? $= I^2; P^2; T^{1-2}$. Dodd's system may be of interest to the pure mathematician, and will probably be of great interest to symbolic logicians whose studies concern the unity of science. Its value to ordinary statisticians and to empirical social scientists will depend upon the applications they can make of this new system of classification. The intellectual history of civilization is full of systems of classification. A new system is useful when it brings isolated and "unknown" things into relationship with some larger synthesis by discovering that these things, first thought to be unique, are in certain respects similar to known things. Perhaps the utility of a new system can be tested by asking: Does it help explain causal relationships? (Dodd claims that his does.) Does it facilitate prediction? Does it facilitate control? For example, Dodd's treatment of control is chiefly in terms of "social control." He does not seem to exploit fully the possibility of classification and sub-classification as a device to promote control (methodological) by providing new or more precise calibrations for matching in experimental designs. Herein lies, perhaps, a fruitful application of his system in empirical social science. No general answer can be made to these questions. The reader can find the answer himself when applied to his own specialty.

F. Stuart Chapin

University of Minnesota

Social Research: A Study in Methods of Gathering Data, by George A. Lundberg. New York: Longmans, Green and Company. 1942. xx, 426 pp. $3.25.

This is the first revision of a book which was published in 1929. It is almost completely rewritten, though it follows the general outline of the original edition in dealing with the philosophy or logic of science and concrete

procedures. Two new chapters are concerned with questionnaires and sociometrics. Other chapters discuss the sample, schedule, measurement of attitudes, opinions and institutional behavior, field work, and social bookkeeping.

Like its predecessor, the revision is on the whole a simple and well-organized survey of problems and difficulties in social research. This comment refers particularly to the section from Chapters V and on in which each chapter contains well chosen examples of techniques of investigation and is concluded with a concise summary and bibliography. As a text, the book should be as successful as the first edition.

The author does not fail to point out both in his introductory four chapters and in critical comments in the summary of special techniques the pluralism of social science in its philosophies of research. If the book has any one distinctive lack, it is the failure to describe this conflict as succinctly and clearly as the particular deficiencies which may occur in sampling, schedule making, and scale construction. Most graduate students will have more difficulty in understanding this lack of coherence in science as such than in the interpretation and use of the instruments of research.

Social statisticians will probably find Chapters VIII, IX, and X to be the most interesting. Therein the explorations of the sociologist in measurement and sociography are described. Similarly Chapter XI on field work should prove to be an excellent guide to a variety of professional students. Chapters VIII-X are complete and constitute as fine an analysis of the subjects discussed as can be found in the literature.

Specific shortcomings in the opinion of the reviewer, in addition to the politics of research, are the omission of a definite statement of what research and science are so that the layman may detect the numerous fads now current under the label of science, too few examples of research methodology from other social science fields than sociology, and failure to stress the point which is mentioned but not amply covered when "judgment" outweighs the virtue of all techniques. Moreover, in the use of studies as examples of research which appear to have unequal scientific merit, there is apt to be some confusion as to the precise limits of what is acceptable as science in sociology.

HAROLD A. PHELPS

University of Pittsburgh

Methods of Correlation Analysis, by Mordecai Ezekiel. (Second edition.) New York: John Wiley and Sons, Inc. 1941, xix, 531 pp. $5.00.

This book contains the same leisurely, readable development of correlation methods which made the first edition an outstanding success. All of the desirable features have been retained and some important improvements in exposition have been made. Of the relatively few additions to subject matter the most important is the emphasis on logical limitations to flexi-

bility of regression functions. Another interesting addition is the study of the reliability of an individual forecast in Chapter 19.

Since it is unnecessary to enumerate the features of a book so well known to readers of this Journal discussion of some changes the reviewer would have liked to see is in order.

None of the recent methods for computing correlation constants based on determinants is seriously considered by the author. In the reviewer's opinion this is an unfortunate omission. Presentation of Stephan's treatment of the wheat problem (this Journal, March, 1931, p. 58) would serve to make a valuable but neglected device available to a wider group of readers.

The use of degrees of freedom in the table of t (Table A) would eliminate confusion in testing correlation and regression coefficients. A different symbol should be used for the number of variables when n is used in the same equation for the number of observations in the sample (p. 211). In view of the great emphasis on "corrected" measures of correlation and "unbiased estimates" of standard errors of estimate a correct statement of the known properties of such estimates is badly needed. They are not median estimates (p. 143); neither are they unbiased in the usual sense of that word. Although discussion of part correlation (involving one independent variable) has been dropped from the text, it is defined in a footnote and derived in Appendix 2. Part correlation involving two independent variables is discussed in squared form (p. 218). The corresponding function appropriate for tests of significance, $(R_{1.23}^2-r_{12}^2)/(1-r_{12}^2)$, is simpler and has a "Beta" distribution with 2 and $n-4$ degrees of freedom.

The author refers to complicated techniques and tests of significance in a manner suggesting a degree of exactness and applicability which they do not possess (pp. 99, 320, 323, 367). The standard error of the mean which is exact under less restricted conditions than any other sampling formula in the book is "proved" in such a way as to suggest that it is approximate. On the other hand, the standard error of r is said to be "precise" (p. 318). The appropriate formula (in which $n-1$ is substituted for $n-2$) is exact for uncorrelated normal bivariate universes, but even the appropriate formula is an approximation for correlated universes. Standard errors for the index of correlation and the multiple correlation coefficient (given without qualification) are greatly in error and of questionable usefulness. The author even advocates the t-test for each of these inherently positive statistics and for "more exact interpretations" he suggests Fisher's z-transformation!

The foregoing criticisms of the material on sampling are not intended to detract from the merits of a great book. According to the preface the book is intended to cover "that portion of the field which is concerned with studying the relationships between variables." Appraised on this basis the book has no equal nor even a close rival.

John H. Smith

Bureau of Labor Statistics

Principles of Punch-Card Machine Operation, by Harry P. Hartkemeier. New York: Thomas Y. Crowell Company. 1942. xiv, 269 pp. $3.25.

While there have been several studies which described specific applications of punched card equipment in accounting and statistical research, the reviewer is not aware of the publication, prior to Dr. Hartkemeier's monograph, of any text purporting to present an elementary treatment of the general problem of machine tabulation. Certainly the manuals which accompany the equipment are entirely inadequate for instructional purposes.

Professor Hartkemeier's text confines itself to instruction in the use of *International Business Machines* equipment and begins with a brief account of the historic development of the Hollerith invention. A statement of the sorting and tabulating principle is followed by a detailed description of the operation of the almost obsolete Type 3 Printer. Next, the standard IBM numeric tabulators (Models 285 and 297) are discussed in great detail. Instruction is given in wiring the tabulator plug board for listing, addition, subtraction, the use of balance counters, class and field selection, and the *X*-distributor.

Out of the 153 pages devoted to the introduction and to numeric equipment, approximately 30 are actually textual; the rest consist of photostat copies of wiring diagrams and of the reports prepared on the tabulator from the various wiring diagrams displayed. Perhaps another 15 pages of text are devoted to the description, in part III, of the alphabetic accounting machine. This more modern apparatus, which the reviewer believes ought to be the almost exclusive subject of a text on the use of punched card equipment, is treated not as the only tabulator with which the modern student of punched card methods is likely to come in contact but rather as an incidental stage in the historic development of the numeric equipment.

Most of the ordinary wiring problems are well discussed and there is a good section on the use of punched cards in obtaining sums of squares and cross products.

While Dr. Hartkemeier has performed a real service in meeting a long felt need for a text on punched card machinery, several possibilities for improvement suggest themselves. The types of IBM equipment discussed constitute a small proportion of that company's offerings in the punched card field. The summary punch, the reproducer, the gang punch, the collator, the multiplying punch, the interpreter, the various models of hand-operated numeric and alphabetic punches, and the mark-sensing reproducer receive little or no mention in this text.

Since no discussion is given of competing equipment, there is no opportunity to point out the types of applications in which rival lines have a material advantage, such as the punching of cards for which a portion of the material is to be repeated from card to card or the listing of material requiring alphabetic characters throughout the line of printing.

It would have been helpful if more attention had been paid to possible fields of application of the techniques described. Except for the section on

progressive digiting (for the accumulation of sums of squares and products) the illustrations are confined to elementary accounting problems.

The reference system by which diagrams and forms are located is not well conceived. The reader may find it confusing to locate diagram 12 on page 61, diagram 12–405 on page 189, and diagram 13 on page 62. The small figures which show the position of the pin in the clearing collar (on page 12) were omitted in earlier printings but this has been corrected in later runs. While the loose-leaf format simplifies the extraction of an individual wiring diagram for comparison with the relevant text and also makes it easier for the student to turn in the blank diagrams supplied with the text after completing them, a number of copies have appeared with defective rings in the loose-leaf binder, which have made the use of the volume somewhat awkward.

Francis McIntyre

Office of Lend-Lease Administration

The Fundamental Principles of Mathematical Statistics, by Hugh H. Wolfenden. New York: The Actuarial Society of America. 1942. xv, 379 pp.

The full title of the book is *The Fundamental Principles of Mathematical Statistics with special reference to the requirements of actuaries and vital statisticians, and an outline of a course in graduation.* The author is a fellow in three of the principal actuarial organizations of Great Britain and the United States and of the Royal Statistical Society. The subject is developed in terms of the needs of a particular field, in this case actuarial work and to a slight extent vital statistics, as contrasted, say, with business, agriculture, biology, psychology or engineering. The book could be described in the words which Yule applied to his own text, "definitely founded on experience, personal experience in statistical work and personal experience in teaching."

There are biographical and bibliographical references throughout the text and a section of 25 pages devoted exclusively to the history of the developments of mathematical statistics. The author explains in the preface that "in teaching these matters there is, inevitably, a cultural responsibilty [beyond] . . . a recital merely of the present state of knowledge. . . . I believe it essential to the proper understanding of any subject to absorb the history of the mental processes which have guided its development." The American reader will note with interest the highly appreciative references to the work by Erastus L. DeForest whose noteworthy and long neglected researches in graduation appeared in *The Analyst* of Des Moines, Iowa, sixty years ago.

Some idea of the scope of the book as compared with typical American textbooks can be gained by a list of exclusions and inclusions. Not found are such topics as kurtosis, time series, index numbers, correlation. The fitting of regression lines and the analysis of variance receive one paragraph each, consisting chiefly of bibliographical references. The meaning of probability is relegated to the second section of the book since the actuarial students

"come . . . with a sound practical knowledge of the elements of the theory of probability." The notation of mathematical expectation is excluded though not the term.

On the other hand, frequency distributions and associated problems of sampling receive extensive treatment; brief but systematic presentation is given the systems associated with the names of Pearson, Gram-Charlier and Poisson-Charlier, and considerable space is given to graduation and fitting of curves for the quantities playing a role in the actuary's office, l_x, $\log l_x$, q_x, and ten other similar functions.

The mean of Bernoulli distribution is found as $m = np$, but with the Poisson distribution, p. 60, the mean is given as $m = nq$. Is it not a "cultural responsibility" to let p be the probability of death by a Bortkewiczian horse kick? The actuarial student can then be told that he will often take the numerical value of p from the q_x column of the life table.

The formula for the standard error of a linear compound, p. 25, could have been found more easily and with greater generality by merely assuming independence, not normality.[1] The student will need further help at p. 30 where variance seems to be referred to as a linear function of the variates, at p. 41 where clarity would come from the concept of all possible samples rather than "n-samples," at p. 42 where the formulae for standard errors of various parameters need such annotations as: mean, "exact"; median, "only approximate even for normal distributions with n large"; standard deviation, "good approximation for normal distributions"; variance, "exact for normal provided the formula is changed from n to $n-1$"; etc. He will need help at p. 52 where the test of the sigmas is presented as "subject to the reservation that ${}_1\sigma_s$ and ${}_2\sigma_s$ must not be unusual," whereas in fact the restriction applies only to the sigma to be used in the denominator.

These and other negative points are offset by insights and generalizations along the way which arouse the reader's gratitude for a thoughtful treatment of the subject from an important point of view at the hands of one who takes teaching seriously and who speaks from wide knowledge and experience.

SIDNEY W. WILCOX

U. S. Bureau of Labor Statistics

National Income and its Composition, 1919–1938, by Simon Kuznets. New York: National Bureau of Economic Research. 1941. Two vols., xxx, 980 pp. $5.00.

This double volume is a report on a major investigation of the country's national income during the indicated years. In Part I, the design and rationale of national income estimation are examined and the author's many elections in choosing between alternatives reviewed. Next, the estimates are presented and analyzed at some length, in terms of totals and different types

[1] See, for example, J. H. Smith, *Tests of Significance*, University of Chicago Press, p. 57.

of components. Part III discusses the derivation of the estimates and notes important characteristics resulting from the nature of concepts, methods, and source materials. This section also compares the estimates with other income computations and introduces the innovation of subjecting the figures to comprehensive tests of reliability. Final sections cover data, sources, and methods in detail. Where pertinent, chapter summaries supply a concise recapitulation of subject matter.

This publication is important to "consumers" as well as "producers" of national income data because of the breadth of vision with which issues are discussed and because of the systematic purposefulness with which the author's investigation of national income was conducted and is detailed to the reader. Naturally, with half its pages devoted to methodology and data it will not normally be studied from cover to cover. Yet the constructive analytic spirit displayed throughout weaves around this type of material an exposition that is remarkable for the insight it provides. Noteworthy is the careful choice made between a sterile counsel of perfection on one hand and ill-considered superficiality on the other—a stimulating example to persons interested in national income.

In view of its scope, this work has remarkably few controversial features. Probably the most important controversy regarding design centers about the use of taxes in valuing governmental services, a choice between "two evils" (p. 32) which leads to the measurement of public services to businesses by the payments of these enterprises to government. Analysis of data tends to be mechanistic. Thus, the study of trends is unduly influenced by the use of averages for certain 5 and 10 year periods because insufficient recognition is accorded the cyclical settings of these averages. The data also have their practical limitations in that they end in 1938, fail to incorporate recently available information, are difficult to use because of the number of variants, and contain intrinsic deficiencies. The last is illustrated by Variant II of the implicit consumers' outlay price index (p. 145) which moves questionably closely with wholesale prices after 1933. But these are relatively minor to this important treatise on a most timely topic.

Dwight B. Yntema

U. S. Department of Commerce

Exchange Control in Central Europe, by Howard S. Ellis. Cambridge: Harvard University Press. 1941. xiv, 413 pp. $4.00.

Professor Ellis has written what is so far the most detailed history of the totalitarian type of control in international trade, leading through the jungle of administrative make-shifts and their economic implications in three countries: Austria, Hungary, and Germany. Extensive analyses in the light (or language) of current monetary theories complete this work of remarkable erudition.

The greater one's respect for the author's painstaking and penetrating analysis, the more disappointing is the interpretative accomplishment. Its

main thesis is that upholding the (nominal) parity with the aid of exchange control was excusable in the crisis, but became harmful, and devaluation should have been substituted in 1930, if not in 1933. Of course, Dr. Ellis knows that *devisen* control was due to extraordinary circumstances, economic and political, but insists on condemning exchange regulations in comparison to open devaluation. He holds Austria as a shining example because she abandoned the regulations (except agreements with clearing countries). But Austria's short term debts were virtually wiped out in the Credit-Anstalt arrangement, her long term foreign debts were effectively reduced, and fresh credit came through the League. Moreover, what the author forgets: Austria received fairly favorable commercial treatment and was encouraged in many ways while the others were to some extent boycotted.[1] Her balance of payments had been brought into sufficient equilibrium to eliminate capital flight, while the other two had to face the "music" throughout the period.

Much of the analysis is scientifically worthless, because it fails to take well-known facts into account. Dr. Ellis imputes repeatedly the vanishing of export surpluses to currency over-valuation, but admits in other places that the currencies were *de facto* devalued. The question, whether bilateralism wasn't the way permitting Germany to maintain even a limited volume of exports and to service partially her debt, is omitted altogether. What about the rapidly increasing foreign resistance against Germany's exports, the exhaustion of her raw material inventories, the effects of a growing internal market due to "reflation" policies, etc.? How can one judge exchange policies without allowing for such factors? Could devaluation protect a debtor against capital flight and the run of foreign creditors? Would not German devaluation have caused repercussions, which an unimportant currency like the Austrian shilling did not provoke? Is it permissible to impute all consequences to a single set of causes when a multitude of (unanalyzed!) factors was at play?

Dr. Ellis believes in the highly controversial Casselian purchasing power parities. Admitting that it is impossible to figure out equilibrium exchange rates under altogether artificial conditions, he proceeds to apply the construct for judging practical policies, using only visible imports and exports at that. The Germans and Hungarians are blamed for not having acted upon an equilibrium rate that existed only in theoretical imagination. The apparent equilibrium rates depended in reality upon expectations of governmental action. Much perspicacity is devoted to comparing German and British "terms of trade." But this concept is meaningless when the economic alternatives are "to be or not to be," to say nothing of political issues. Lastly, Dr. Ellis operates under a twofold misunderstanding: that exchange control necessarily leads to autarchy and totalitarianism; and that its abolition means the return to normal trade and political systems. In reality, exchange control was merely the most extreme among the super-protectionist systems

[1] Neither the importance of tourist traffic for Austria's balance of payments, nor the preference given Austria by Western tourists is mentioned.

of the 30's; the same objectives could be, and have been, accomplished in the framework of the British and French types of exchange policies. The alleged results of bilateralism—reducing the volume of trade, changing its direction and composition, and extending monopolistic control—were obtained elsewhere by multilateral methods. The greatest weakness of the book is the lack of a comparative picture, and the consequent bias against one system.[2]

Melchior Palyi

Chicago

The Federal Reserve Bank of Cleveland, by Arthur F. Blaser, Jr. New York: Columbia University Press. 1942. xxvii, 300 pp. $3.50.

The Federal Reserve Bank of Richmond, by Charles G. Coit. New York: Columbia University Press. 1941. xv, 140 pp. $2.00.

The organization of these books and the questions with which they deal bespeak their common origin. They are the most recent of a series of studies initiated by the Banking Seminar at Columbia University to cover each of the twelve Federal Reserve Banks. Others have already been published for the New York, Chicago, San Francisco and Boston Reserve banks.

Well organized, brief, and readable, Mr. Blaser's and Mr. Coit's studies describe the commercial, industrial, agricultural, and financial activities and institutions of the Cleveland and Richmond Federal Reserve districts. The authors' main question is the extent to which the Cleveland and Richmond Reserve banks have operated as semi-autonomous central banks in their respective districts, as intended by the framers of the original Federal Reserve Act. They mutually conclude that the banks are rather administrative branches than semi-autonomous units. To reach this conclusion hardly requires a couple of books. It will go unchallenged. The same conclusion has been reached independently by many students, and by more persons who are not, who are interested in banking and aware of the obvious tendencies of government in the better part of the past decade.

The one possible exception in which the authors concede some autonomy of local action is discounts, since these are made directly by the Reserve banks. But, as the authors point out, there have been no discounts worth mentioning for some years, and discount policy, including the rate, is subject to Reserve Board regulation.

One achievement, however, can be credited to the framers of the Federal Reserve Act in the matter of local autonomy. Thanks to them, the Reserve banks have been staffed for the most part with men who have originally come from the districts or have become established residents there. This gives to bankers and local business men a measure of assurance that they are dealing with people who have some knowledge and sympathy for their own

[2] Dr. Ellis' bibliography indicates his bias by omitting such pertinent literature as J. Trier, H. Luckas, F. Sohaum, P. Haber, Perfot, etc. Serious students of the 30's, like H. J. Tasca, K. E. Poole, H. Heuser, and C. W. Guillebaud, who emphasized the difficulties of debtor countries, are barely noticed or actually ridiculed, while a dilettantic propagandist like Th. Balogh is approvingly quoted.

conditions and problems. Thus, in the realm of administration, if not in the realm of policy, a large measure of local autonomy has been achieved, which the Federal Reserve Governors on more than one recent occasion have pointed out as administratively desirable.

VICTOR M. LONGSTREET

Board of Governors of the
Federal Reserve System

Federal Crop Insurance in Operation, by J. C. Clendenin. Stanford University: Food Research Institute. Wheat Studies, Vol. XVIII, No. 6. March 1942. 229–290 pp. $1.25.

This paper outlines extensively and discusses many phases of the Federal Wheat Crop Insurance Program. It reports the results of study of a wide variety of information such as regulations and procedures, basic individual farm data, questionnaire surveys sent to farmers and businessmen, conversations with farmers and with planning, administrative and operational employees. The insurance contract, actuarial details, participation, unit costs, public relations, administration and finances also are treated. Footnotes give many details. There is an appendix of state and county data and a brief international history of crop insurance. The report is typical of those sponsored by the Food Research Institute in breadth, clearness and balance.

The annual contract guarantees a yield equal to 75 per cent (or 50 per cent at the election of the applicant) of average for the farm. The Corporation collects premiums and pays indemnities in either the wheat or the cash equivalent of a specified class and grade but does not protect against loss in quality. The per acre cost of insurance closely approximates the annual average bushels per acre loss which would occur on the farm in a representative period of past years. The individual farm rating basis rather than a community or county basis is approved but it is suggested that even a smaller or individual field unit rating basis may become advisable.

The guarantee of yield rather than value is commended as economically sound. The scope and interpretation of insurance coverage is called "quite satisfactory." The risks covered include practically every hazard beyond the control of the insured and the risks not covered are those which are due to neglect, malfeasance, theft, fraud, or unsound farming practices. Loss adjustments are "reputed to be fair, sometimes verging on severe, and to reveal few cases of lavish settlements." The author concludes that this is encouraging for laxity would quickly demoralize the program. A suggestion is made, however, that there is some probability that adjustment standards may be responsible for part of the underwriting losses.

Excess of indemnities paid over premiums collected in each of three years is attributed also to inaccurate and incomplete basic farm data and method of use, absence of any but casual check on the fields seeded except for acreage, and unsatisfactory distribution of risks.

The Federal Crop Insurance Corporation's program is coordinated with

that of the Agricultural Adjustment Agency. Mention is made of the low cost of field administration made possible by the use of county AAA committees. The value of their knowledge of local conditions and of their sales prestige is properly emphasized. However, the "caliber" and attitudes of the committees vary. Some are not insurance minded or are untrained as salesmen and, in part, are responsible for uneven distribution of risks and lax appraisals.

The extent and volume of participation, which cover about one-fifth of the text, are called "phenomenal" in the first two years. "The third year brought a sharp decline in the rate of growth . . . and yet policies covered over 17 per cent of the total seeded acreage and about the same percentage of the nation's wheat growers." The author seems convinced that the distribution of business shifts chiefly with shifts in soil moisture reserves and recent loss experience but he proceeds inconsistently to stress the importance of farmer sales resistance to relatively high rates in the high-risk area.

A public policy section of the paper expresses doubt that insurance encourages uneconomic land use, discusses premium wheat storage problems, and considers the adaptability of the program to types of farming regions.

The Federal Government bears the administrative cost of the program. Unit costs are thought to be high and in the main irreducible except by a material increase in business volume, but this opinion is based on several questionable assumptions.

The author concludes finally that wheat crop insurance should be given a thorough trial by testing several alternative details before it becomes crystallized. A schedule rate, premium reduction for small loss experience, and a term policy are important alternatives suggested. Similar plans have been under consideration for some time by the Corporation and are now a part of the contract being currently offered.

Such a well-rounded picture of the activities of the Corporation under one cover has been lacking in the published literature. There are some inaccuracies and questionable assumptions but they are not important insofar as a comprehensive presentation of the program is concerned. The paper merits a well accepted contribution.

Richard O. Cromwell

Federal Crop Insurance Corporation

American Highway Policy, by Charles L. Dearing. Washington: The Brookings Institution. 1941. xi, 286 pp. $3.00.

The public policy problems involved in building, maintaining, financing, and policing our highway system of more than three million miles are obviously numerous and complex. These problems though interrelated may be conveniently grouped in three broad categories: (1) determining a satisfactory method of distributing authority and responsibility among the several levels of governments; (2) allocating the financial burden among individual taxpayers; and (3) resolving the long-standing controversies

between highway transportation enterprises and competing modes of transportation. The present volume was written in response to a request by the Commissioner of Public Roads of the United States Public Roads Administration addressed to the Brookings Institution to institute a study of the underlying principles according to which the abundant data already collected by the state-wide planning surveys might be employed in the solution of the controversial problems of highway financing and administration.

Mr. Dearing observes that the provision of public roads has always been deemed to be an essential service of government. He states that "the modern road plant is a multiple-purpose facility," serving "in one degree or another to give access to land and buildings; to facilitate the movement of goods and people primarily associated with community life; to supply the avenues of optimum intercommunity mobility; and ... to expedite the administration of various essential functions of government." The author proposes to have the roads in each state classified along lines of use, with the so-called "general purpose" roads under the exclusive jurisdiction of the state agencies. He would assign the major financial burden of supporting the general purpose roads to motor vehicle users under "a system of levies designed to measure differential road occupancy as well as the cost of any additional physical facilities provided to meet the requirements of unusual size and weight characteristics." The remaining roads, which he says are generally used for community service and land access purposes, would be administered and mostly financed by local units of government.

Mr. Dearing believes that Federal participation in the road program should be confined to "those types of activities which are designed to serve broad national objectives." In order to help limit the activities of the Federal Government in the highway field, he proposes to reserve to the states exclusive jurisdiction over the administration of the special motor vehicle user charges. He also expresses strong opposition to the so-called "public utility method" of managing the entire road plant as if it were a regulated public utility competing at all points with the railroads.

In this limited space, there is opportunity only to suggest, by presenting a few examples of his views, that the author has identified and thoughtfully analyzed the salient issues in highway public policy. The book is well organized and concisely written and will repay close study by all who are concerned with highway problems.

RALPH L. DEWEY

Washington, D. C.

The Marketing of Used Automobiles, by Theodore H. Smith. Columbus: Bureau of Business Research, Ohio State University. 1941. xv, 200 pp. $3.00.

This book covers a larger field than is suggested by its title and is really a study of the marketing of both new and used cars from the beginning of automobile production through the year 1940. The reason for the inclusion

of new car sales information is that the two subjects are virtually inseparable. Used cars are, in the main, sold by dealers in new cars, while used car prices and sales volume are directly and powerfully affected by new car prices and sales volume.

A great amount of factual information is presented, much of it in statistical form, including forty-five tables and eleven charts. The author's comments and explanations are lucid and intelligent. The book should be a convenient source of reference for related information, such as the origin and growth of the sales finance companies, and various sidelights on the motor vehicle manufacturing business. For instance, we are told that the *National Used Car Market Report* for June 1915 gave price information on 154 makes of gasoline and steam cars and 14 makes of electric cars, these being "only the better known makes." This is quite a contrast to conditions in 1941, when there were only 12 manufacturers of passenger automobiles in the United States.

The book is not, as might be supposed, a compendium of advice to dealers on how to sell used cars, but is rather an historical treatise on how used cars have been sold, with discussion of the practices used, and their results. However, it would doubtless be worthwhile for any erstwhile automobile dealers who expect again to be in that business, when it is restored to life after the war, to read the book.

The automobile business differs from all others in the following respects: (1) The new product is sold to the public only through "enfranchised" dealers; (2) these dealers are required to pay in full for their stock in trade before delivery; (3) in the majority of sales the dealer has to accept used merchandise as part payment of the price of the goods sold; and (4) due to the high unit price, the majority of buyers cannot raise enough money to pay in full at time of purchase, but must buy on instalments, or not at all.

These conditions produce, among others, the following results: (1) The manufacturers have the power, which they have always used, more or less, to "coerce" the dealers as to methods of conducting their business; for the manufacturer can cancel the dealer's franchise, and put him out of business; (2) the average dealer cannot exist without wholesale and retail financing. He must borrow from a finance company most of the money needed to buy his new cars and he must sell to a finance company the instalment contracts resulting from his retail sales; (3) sales competition between dealers centers about the trade-in allowance rather than the retail sales price; (4) the dealer almost always loses money on the trade-in. He allows more for it than he can sell it for.

It is because of these conditions that many unique problems arise in the retailing of automobiles and especially of used automobiles. Mr. Smith's book discusses all of these problems in a highly intelligent manner.

Milan V. Ayres

Chicago

The Development of American Industries, planned and edited by John G. Glover and William B. Cornell. New York: Prentice-Hall, Inc., 1941. xxviii, 1005 pp. $5.50 (trade) and $4.50 (school).

This revised edition follows closely the first edition issued in 1932. In the main the exact content is retained except for modifications to incorporate more recent data and to cover new legislation affecting industry, new methods of production, and new products. Some of the data, however, remain obsolete or are otherwise not fully satisfactory for measuring conditions in the industries or changes therein. Specifically, where reliance has been placed on data for scattered years, the measurement frequently reflects temporary conditions more largely than basic trends. This weakness could have been overcome by a more careful selection of comparable periods or by the introduction of charts to present a more complete picture in the available space.

Although the authors state in the preface that "the effect of war conditions abroad and of our own national defense program has not been unduly emphasized," they would have been more nearly correct to state that they have given almost no consideration to that phase. This is true even though the metal, machine-tool, aircraft, power, chemical, and shipbuilding industries, covered in various chapters, were vitally affected long before the publication of this edition in September 1941.

A survey of prewar industries, of course, may provide a useful background for the study of current problems. This work remains one of the best collections of essays for the non-technical reader on industrial development. Altogether it contains 39 chapters on 38 industries, plus an initial chapter on labor's contribution to American industries and a concluding chapter on trade associations. Most of the chapters have been prepared by executives in business concerns or trade associations. The contributions are of highly variable quality. Owing to the sparsity of references, the intellectually curious reader will, no doubt, be left with a feeling that he has been given a hasty introduction without additional guidance. If he chooses to investigate, he will find that many of the statements are drawn from readily available sources. References would aid also in checking on information, which is not always reliably presented. Several summaries of legislation, for example, are misleading to the extent of confusing proposals with actual enactments. This deficiency of references and directives to outside sources is counterbalanced to some extent by an unusually good index to the material within the volume.

WILBERT G. FRITZ

National Resources Planning Board

PUBLICATIONS RECEIVED*

Adams, George P., Jr. *Wartime Price Control.* Washington: American Council on Public Affairs. 1942. x, 153 pp. $3.00 (cloth), $2.50 (paper).

American Public Health Association, The Committee on the Hygiene of Housing of the. *Housing for Health.* Lancaster, Pennsylvania: The Science Press Printing Company. 1941. 221 pp. $1.00.

Black, John D. *Parity, Parity, Parity.* Cambridge: The Harvard Committee on Research in the Social Sciences. 1942. xi, 367 pp.

Blaisdell, Donald C. *Government Under Pressure.* New York: Public Affairs Committee, Inc. Public Affairs Pamphlets No. 67. 31 pp. 10 cents.

Bliss, Charles A. *Operating Results of Department and Specialty Stores in 1941.* Boston: Harvard University Graduate School of Business Administration, Bureau of Business Research. 1942. vi, 62 pp. $2.50.

Brewington, Ann, and Berg, Evelyn Van Emden. *The Women Graduates of a Collegiate School of Business.* Chicago: The University of Chicago Press. The School of Business. Studies in Business Administration, Vol. XII, No. 2. 1942. ix, 60 pp. $1.00.

Buechel, F. A., assisted by Dedeke, Edward R. *Directory of Texas Wholesale Firms.* Austin: The University of Texas, Bureau of Business Research. 1942. 380 pp. $2.00.

Dean, Joel, and James, R. Warren. *The Long-Run Behavior of Costs in a Chain of Shoe Stores.* Chicago: The University of Chicago Press. The School of Business. Studies in Business Administration, Vol. XII, No. 3. 1942. ix, 54 pp. $1.00.

Delaware, University of. *Delaware Income Statistics Compiled from Income Tax for 1936, 1937, and 1938.* Newark: University of Delaware, Bureau of Economic and Business Research. Volume I. 1941. xliii, 107 pp.

Ford, Robert S., and Goodrich, Kenneth. *State Supervision of Local Borrowing.* Ann Arbor: University of Michigan. Bureau of Government. 1942. 41 pp. 10 cents.

Ford, Robert S., and Waxman, Albert. *Financing Government in Michigan.* Ann Arbor: Bureau of Government, University of Michigan. Michigan Governmental Studies, No. 9. 1942. vi, 213 pp. $1.00.

Freeman, H. A. *Industrial Statistics.* New York: John Wiley & Sons, Inc. 1942. ix, 178 pp. $2.50.

Galloway, George B. *Postwar Planning in the United States.* New York: The Twentieth Century Fund. 1942. xi, 158 pp. 60 cents.

Gault, Edgar H., and Goodman, Charles S. *The Control of Customer Returns.* Ann Arbor: University of Michigan, School of Business Administration, Bureau of Business Research. 1942. 107 pp. $1.50.

Haberler, Gottfried. *Consumer Instalment Credit and Economic Fluctuations.* New York: National Bureau of Economic Research, Inc. 1942. xix, 239 pp. $2.50.

Haig, Robert Murray. *Financing Total War.* New York: Columbia University Press. 1942. 32 pp. 25 cents.

Halm, George N. *Monetary Theory.* Philadelphia: The Blakiston Company. 1942. xi, 347 pp. $3.50.

Hargy, H. W., and Mitchell, R. V. *Stock Control Methods.* Urbana: University of Illinois Bureau of Econom-

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. Some items of minor interest to statisticians have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications which are to be found in the *Revue de l'Institut International de Statistique, Journal of the Royal Statistical Society, American Economic Review, Population Index, Transactions of the Actuarial Society of America, The Record of the American Institute of Actuaries,* and *Sankhyā, The Indian Journal of Statistics.*—Editor.

ic and Business Research. Business Studies, No. 1. 1942. 20 pp.

Kerby, C. Edith. *Manual on the Use of the Standard Classification of Causes of Blindness.* New York: The Committee on Statistics of the Blind, jointly sponsored by the American Foundation for the Blind and the National Society for the Prevention of Blindness. 1940. v, 26 pp.

New York City, Board of Education. *Forty-Third Annual Report of the Superintendent of Schools of City of New York.* 372 pp.

Nourse, Edwin G. *Between Automatic and Authoritarian Price-Making.* Washington: The Brookings Institution. Chapter I of the forthcoming book *Price-Making in a Democracy.* 1942. vii, 38 pp. 25 cents.

——— *Free Enterprise and Laissez Faire.* Washington: The Brookings Institution. Chapter II. 1942. vii, 26 pp. 25 cents.

——— *Competition as Method and as Goal.* Washington: The Brookings Institution. Chapter III. 1942. vii, 30 pp. 25 cents.

——— *The Profit Motive and "Maximum" Profits.* Washington: The Brookings Institution. Chapter IV. 1942. vii, 30 pp. 25 cents.

——— *Free Enterprise, Price Policy, and Democracy.* Washington: The Brookings Institution. Chapter V. 1942. vii, 34 pp. 25 cents.

——— *Areas of Applicability of Low-Price Policy.* Washington: The Brookings Institution. Chapter VI. 1942. v, 51 pp. 25 cents.

Orchard, John E. *Resources for Victory.* New York: Columbia University Press. 1942. 30 pp. 25 cents.

Palmer, Gladys L. *The Significance of Employment Patterns in Households for Labor Market Analysis.* Philadelphia: Wharton School of Finance and Commerce, University of Pennsylvania, Industrial Research Department. Special Report No. 8. 1942. 21 pp.

Schumacher, F. C., and Chapman, R. A. *Sampling Methods in Forestry and Range Management.* Durham: Duke University School of Forestry, Bulletin 7. 1942. 213 pp. $2.50 (cloth), $2.00 (paper).

Simutis, Anicetas. *The Economic Reconstruction of Lithuania After 1918.* New York: Columbia University Press. 1942. xiii, 148 pp. $1.50.

Special Libraries Association. *United States Government Periodic Publications.* New York: Special Libraries Association. Compiled by the Washington, D. C., Chapter. 1942. v, 87 pp. $2.00.

Spengler, Joseph J. *French Predecessors of Malthus.* Durham: Duke University Press. 1942. ix, 398 pp. $4.50.

Stein, Emanuel, and Backman, Jules, Editors. *War Economics.* New York: Farrar & Rinehart, Inc. 1942. x, 601 pp.

Wald, Abraham. *On the Principles of Statistical Inference.* Notre Dame: University of Notre Dame. 1942. 47 pp.

W. K. Kellogg Foundation. *The First Eleven Years, 1930–1941.* Battle Creek, Michigan. 1942. 217 pp.

Government Publications

Department of Agriculture. *Farm-Mortgage Credit Facilities in the United States.* Miscellaneous Publication No. 478. 1942. vii, 262 pp. 55 cents.

——— *An Objective Method of Sampling Wheat Fields to Estimate Production and Quality of Wheat,* by Arnold J. King, Dale E. McCarty, and Miles McPeek. Technical Bulletin No. 814. 1942. 87 pp. 10 cents.

——— The Bureau of Home Economics (*in cooperation with the Work Projects Administration*). Consumer Purchases Study, Urban and Village Series. *Family Expenditures for Clothing, Five Regions.* Miscellaneous Publication No. 422. iii. 320 pp. 35 cents.

——— *Family Expenditures for Personal Care, Gifts, Selected Taxes, and Miscellaneous Items, Five Regions.* Miscellaneous Publication No. 455. iii, 113 pp. 20 cents.

——— *Family Expenditures for Housing and Household Operation, Five

Regions. Miscellaneous Publication No. 457. v, 201 pp. 25 cents.

——— *Family Income and Expenditures, Southeast Region.* Part 1, Family Income. Miscellaneous Publication No. 462. v, 208 pp. 25 cents.

——— *Changes in Assets and Liabilities of Families, Five Regions.* Miscellaneous Publication No. 464. v, 220 pp. 25 cents.

——— *Family Income and Expenditures, Five Regions.* Part 2, Family Expenditures. Miscellaneous Publication No. 465. v, 366 pp. 35 cents.

Department of Commerce, Bureau of the Census. *Admissions to the Baltimore City and the District of Columbia Jails: 1940.* 1942. 33 pp.

——— Bureau of Foreign and Domestic Commerce. *Trade and Professional Associations of the United States,* by C. J. Judkins. Industrial Series No. 3. 1942. vi, 324 pp. 70 cents.

Federal Deposit Insurance Corporation. *Assets and Liabilities of Operating Insured Banks, December 31, 1941.* Report No. 16. 1942. 35 pp.

National Archives. *Seventh Annual Report of the Archivist of the United States, 1940 1941.* Publication No. 20. 1942. vii, 95 pp.

National Resources Planning Board. *National Resources Development Report for 1942.* 1942. 227 pp. 55 cents.

——— *Our Public Works Experience.* June 1941. 36 pp. 10 cents.

Department of State. *Proceedings of the Eighth American Scientific Congress.* Held in Washington May 10–18, 1940. Volume I (English Edition). Organization, Activities, Resolutions, and Delegations. 539 pp.

Alabama State Board of Health. *Report Relating to the Registration of Births, Stillbirths, Deaths, Marriages and Divorces in the State of Alabama for the Year Ending 31st December 1940.* Montgomery. 1941. clxvii, 301 pp.

Tennessee Department of Public Health. *Annual Bulletin of Morbidity Statistics Tennessee 1941.* Nashville. 1942. 54 pp.

Advertising Service Guild, The. *Home Propaganda,* A Report prepared by Mass-Observation for the Advertising Service Guild. London: The Advertising Service Guild. *Change,* No. 2. 1942. vi, 78 pp. One shilling.

British Information Services. *The Taxation of Weekly Wage Earners.* London: His Majesty's Stationery Office. Presented to Parliament by the Financial Secretary to the Treasury by Command of His Majesty. 1942. 16 pp. 3 d. net.

House of Commons. *First Report from the Select Committee on National Expenditure.* London: His Majesty's Stationery Office. 1942. 48 pp. 9 d. net.

Bureau de Statistique de l'Islande. *Manntal á Íslandi, Recensement de la Population de l'Islande.* Reykjavik: Prentað í Ríkisprentsmiðjunni Gutenberg. Hagskýrslur Íslands 92. 1937. 160 pp. 4 krónur.

——— *Alþingiskosningar, Élections au Parlement 1937.* Reykjavik. Hagskýrslur Íslands 96. 1938. 37 pp. 1 króna.

——— *Mannfjöldaskýrslur, État et Mouvement de la Population 1931–1935.* Reykjavik. Hagskýrslur Íslands 99. 1938. 90 pp. 2 krónur.

——— *Búnaðarskýrslur, Statistique de l'Agriculture en 1939.* Reykjavik Hagskýrslur Íslands 106. 1940. 61 pp. 1 króna og 50 aurar.

——— *Verzlunarskýrslur, Statistique du Commerce en 1939.* Reykjavik. Hagskýrslur Íslands 107. 1941. 139 pp. 2 krónur og 50 aurar.

——— *Fiskiskýrslur og Hlunnindi, Statistique de la Pêche, de la Chasse au Phoques et de l'Oisellerie en 1939.* Reykjavik. Hagskýrslur Íslands 108. 1941. 54 pp. 2 krónur.

República Dominicana, Dirección General de Estadística Nacional. *Anuario Estadístico de la República Dominicana.* Ciudad Trujillo. 1942. 2 Volumes 854 pp. (Vol. I), 883 pp. (Vol. II).

Ministerio de Trabajo, Dirección General de Estadística. *Anuario Estadístico de España.* (Edición manual). Madrid. 1941. xvi, 521 pp.

Eidgenössischen Statistischen Amt (Bureau fédéral de Statistique). *Statistisches Jahrbuch der Schweiz (Annuaire Statistique de la Suisse).* Berne. 1941. ix, 495 pp.

——— *Neue Schweizerische Sterbetafeln.* Berne: Beiträge Zur Schweizerischen Statistik, Heft 10. 1942. 50 pp.

——— *Population de Résidence des Communes. Résultats provisoires.* Berne: Bureau Fédéral de Statistique. Recensement Fédéral de la Population. 1942. 50 pp.

// JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 220 DECEMBER, 1942 *Volume 37*

COST OF LIVING INDEXES IN WARTIME

By Faith M. Williams, Frances R. Rice and Emil D. Schell,
U. S. Bureau of Labor Statistics

The cost of living statistician is always in a difficult situation in a period when consumers' buying is changing rapidly. The theory of a cost of living index is that it measures changes through time in the cost of the goods and services customarily purchased by a given group, in a given period, always pricing goods of the same quality. Changes in consumer buying in 1942 are occurring much more rapidly than in the 1920's and 1930's, but many of the problems faced by the statistician at the present time are similar to those met over the last two decades. In wartime, however, it is always more difficult than in peace time for cost of living statisticians to mind their p's and q's.

In the latter half of the 1920's, Mr. Ethelbert Stewart who was then Commissioner of Labor Statistics began to make plans for a new study of consumption patterns on the ground that consumption habits had changed considerably since 1917–19 (the period to which the weights for the Bureau's cost of living index applied). The funds were not made available, however, until 1934 and it was only then that the basic survey was begun which made it possible to take full account in the Bureau of Labor Statistics' cost of living index of the changes in consumption habits since 1919.

It has been explained elsewhere (see Bureau of Labor Statistics' Bulletin 699) that the actual computation of the index did perforce take account of many changes in consumer purchases from 1919 to 1935. A number of the goods included in the index, when it was first calculated, entirely disappeared from the market long before 1935. High button shoes were succeeded by oxfords; pajamas replaced nightshirts; dining room suites replaced dining table, buffet and chairs sold separately. It would have been impossible to proceed with the calculation of the index if many such substitutions had not been made.

After the Bureau's study of the expenditures of wage earners and clerical workers was made in 1934–36, a new set of weights was prepared for each of the city indexes which are included in the Bureau's index for the large cities of the United States. These new weights were

linked with old weights as of March 1935, and a new series computed with new weights from that date.

The number of goods and services covered in the Bureau's new index which had never before been included in the cost of living index provides a rough indication of the changes which occurred in the purchases of wage earners and clerical workers in the United States between the two periods to which the weights for the Bureau's index apply, 1917–19 and 1934–36. The most important of the additions are as follows: Automobiles, automobile repairs, gasoline, motor oil, tires and tubes, automobile licenses and taxes, automobile insurance, electric refrigerators, gas refrigerators, washing machines, vacuum cleaners, radios, electric light bulbs, felt-base floor coverings, living room suites, railroad fare, silk and rayon yard goods, certain fresh fruits and green vegetables. In addition various services were included for the first time: hospital room, surgeon's fees for appendectomy and tonsillectomy, domestic service, beauty shop services, and dry cleaning.

In the last few months, some goods that Americans have customarily purchased for the last 15 or 20 years have suddenly disappeared from the market for the duration; serious shortages have developed in certain staple goods, and some of these like sugar are now being rationed; important quality changes have appeared in other goods. As the production machinery of the country is converted to war purposes, the goods and services produced for consumers are undergoing more drastic changes in the course of a single year than occurred in the entire period from 1919 to 1936—but in the opposite direction. If an index had been computed to show changes in the plane of living of the people of the country, it would have shown a sharp increase in 1941, and would be showing a sharp decline in 1942. But an index of the actual plane of living is quite different from an index of the cost of living.

The drastic changes which have occurred in the supply of consumers' goods available to civilians since December 1940 have made necessary a number of changes in the computation of the Bureau of Labor Statistics' cost of living index. These changes have occurred so rapidly that it has frequently been necessary to make changes without an adequate statistical base. It has not been possible, nor would it have been desirable, as it was when the Bureau produced a new cost of living index as of March 1935, to link in a new bill of goods and services at one date. Adjustments in the index to the new conditions, therefore, have been made currently.

An alternative procedure would have been to hold constant the price of the goods no longer available and to make no change in the weighting of the index, regardless of whether goods are available. If

this procedure had been adopted, it would have been possible to invalidate the cost of living index as a tool for adjusting wages to changes in living costs on the ground that it had lost its connection with reality, that it no longer measured the cost of goods actually available.

Rationing adjustments. Under the first major rationing order which applied to civilians in the United States, new automobiles were withdrawn from the market for an indefinite period. The second order provided that special permission would be required to purchase new tires and tubes and that permission would be granted only in exceptional cases. In January, 1942, accordingly, the Bureau of Labor Statistics adjusted its cost of living index to this situation. New automobiles and new tires were dropped from the index. Used automobiles and used, recapped, and retreaded tires were added, with weights representing the best available estimates of the limited extent to which they were being purchased by wage earners in January and February. The relative importance of automobile repairs and public transportation costs was increased in computing the index. Judging from reports from the Bureau's field agents, a smaller amount of money was being spent by the moderate-income urban families early in 1942 for transportation of all types, than had been spent formerly. Accordingly, the relative importance of transportation costs was reduced from 29 to 23 per cent of the miscellaneous index, and from 8 per cent to about 6 per cent of all living costs. The aggregate cost of the eliminated goods was transferred from the transportation group to the group of unallocated items for which it is assumed prices move with the aggregate of costs for all goods and services priced for the index. The relative importance of the other groups of goods and services included in the miscellaneous index, therefore, was increased slightly.

With further rationing orders, prices of all tires were dropped from the index in April. In May, in the area where gasoline was rationed, automobiles were dropped entirely and drastic cuts made in the weights for gasoline and motor oil. In the areas where gasoline was not rationed, the weight for automobiles was cut in half and for gasoline and motor oil, cut one-third. In November, when gasoline rationing will be universal throughout the country, a further adjustment in transportation weights will take place.

The relative importance of the several transportation costs to moderate-income workers in large cities, as it has changed in the computation of the cost of living index, is shown in Table I.

Disappearance of goods. As striking, perhaps, as the changes in consumer buying resulting from actual rationing orders have been the changes in purchasing habits caused by the disappearance of articles

the index of a considerable degree of accuracy in the comparison of prices on identical articles from month to month.

Where changes in quality are so concealed that the agents cannot detect them, it has never been possible to make any allowances for quality changes, and will not be possible until programs of quality labeling are developed and put into practice.

Scientific testing of consumers' goods for durability and for efficiency in performing the services for which they are intended is still in its infancy. Performance tests have been developed by the Federal government in order to improve its purchasing procedures, and in order to administer the tariff or acts intended to protect the consumer, or to improve trade practices. Other performance tests have been worked out by industrial engineers working for manufacturers or distributors. In some cases these industry tests have been reviewed by the scientific societies concerned and by the American Standards Association, and have become standard practice; in other cases, as for example the test of heat conductivity for textiles, no test method has received the final approval of any scientific group. As far as the present writers have been able to ascertain, no attempts have been made to prepare measures of efficiency for different types of consumers' goods which would combine several different physical measurements. We have found no statistical comparison of the serviceability, for example, of a specified overcoating made of cotton and reprocessed wool with a specified overcoating made entirely of new wool. The textile chemist can supply test results on resistance to abrasion, breaking-strength, and heat conductivity, although the last-named test is not so well standardized as the others. No one has, however, attempted to work out a formula for weighting the results of these tests into figures which can be used to compare the relative efficiency of the two textiles from the point of view of consumer use.

Today the problem is more acute than ever before. The Bureau's agents must watch not only for quality changes, but also for the changes in sizes and package weights which are reported to have been made since price controls have been imposed. For example, the Bureau's agents noted in 1941 in some stores that a box of corn flakes, priced the same as previously, was marked as containing 6 ounces, where it formerly held 8 ounces. This change appeared in the index of food costs as an increase in the price of a half-pound of corn flakes.

The Bureau follows the policy of continually developing new specifications, as goods formerly purchased are no longer on the market. In pricing for the September, 1942, index, over 100 new specifications

were used, for example, 5 grades of rayon stockings and one grade of cotton anklets. Instructions to the field agents on the use of these specifications, and the specifications for one grade of rayon stockings follows:

> *Rayon hosiery.* Hosiery manufacturers have been consulted as regards these specifications. Cotton reinforcements in the toe, heel, sole and welt are being used when the yarn is available because it is more durable. Viscose yarn has been specified for each item. The use of acetate rayon in hose is still in the experimental stage but will probably become more important in the future.
>
> Hose, rayon, 100 denier, 45 gauge. FABRIC: Viscose rayon, high twist; heel, sole and toe reinforced with cotton, 150 denier rayon welt. CONSTRUCTION AND STYLING: Full fashioned, plain knit welt with run stop, full coursed, not more than 6 flare narrowings.

(Prices are obtained separately for brands widely advertised by manufacturers, and brands not advertised or advertised locally only by distributor or manufacturer.)

Figures on mill shipments of hosiery, retailers' reports, and the observations of the Bureau's field agents indicated that stocks of silk stockings in retail stores were running low. On that account the weight on silk stockings in the Bureau's September index was reduced by two-thirds, and rayon stockings and cotton socks were linked into the index. The weights for rayon stockings and cotton socks were estimated from the latest figures on mill shipments.

The computation of the rent index, again, provides much the same theoretical problems. New dwellings must be linked into the index, when dwellings formerly priced are no longer inhabited, or have deteriorated in quality. However, where decreases have occurred in facilities available, with no change in money rent paid, no increase in rent is recorded.

In the past, all rental quotations obtained by the Bureau for inclusion in the index were obtained from rental agencies. At the present time, however, many property owners are withdrawing rental properties from rental management agencies, so that a sample based on rental management agencies would no longer be representative. In addition, there is some evidence that private owners who handle their own properties are not complying as fully with the ceilings established by the OPA as the real estate agencies who make a regular business of rental management. Beginning with September, 1942, a sample of rent quotations obtained from tenants were collected in 13 of the 34 cities, and after a period of experimentation, were included in the index, in place of the reports from management companies.

Taxes in the index. Excise and sales taxes levied directly upon the price of goods and services sold to consumers have always been included, in the computation of the Bureau's cost of living index, as a part of that cost. Taxes applicable at retail are included by adding the amount of tax to the retail price of each commodity in the index on which such taxes have been imposed. Excise taxes levied at the manufacturers' level are automatically included in the price obtained for inclusion in the index, to the extent that those taxes are passed on to the consumer, since the article is priced in retail stores.

The extent of the price rise due to excise taxes, as distinguished from the extent due to other causes, can be determined, and has been published, on several occasions.[1] This procedure is the same as that followed for the British index. In the July, 1942, issue of the *Ministry of Labour Gazette* (p. 141), there appears the following statement:

> The rise of 45 points since the beginning of September, 1939, is equivalent to about 29 per cent. Of these 45 points, about 4 points represent the effect of the increases, since that date, in the taxes on sugar, tobacco, and cigarettes, and matches; and approximately 6 points are due to increases resulting from the Purchase Tax.

Insofar as taxes on real estate are reflected in higher rental costs, these taxes, too, serve to increase the cost of living, as measured by the Bureau's index. Income taxes, on the other hand, are considered as deductions from income and not as current expenditure.

The amount of the tax to moderate-income families can, of course, be readily computed. In Sweden, this is done, and the change in the amount of the income tax paid by families covered by the cost of living index is published. In addition, the total cost of living index in Sweden, unlike that computed for this country, includes the income tax.

WARTIME ADJUSTMENTS IN THE COST OF LIVING INDEXES OF CANADA, GREAT BRITAIN, AND SWEDEN

The Dominion Bureau of Statistics has, as yet, announced no change with regard to the computation of the Canadian cost of living index, and it may, therefore, be inferred that this index reflects the quality changes in the pre-war plane of living. Few durable goods are priced for inclusion in the Canadian index so that it is unlikely that war time

[1] October 15, 1941, report; BLS #14160 "Changes in Cost of Living in Large Cities."
"The cost of goods purchased by wage earners and lower-salaried workers in large cities increased 1.2 per cent from mid-September to mid-October. Advances in rents and in prices of food, clothing, automobiles, and certain housefurnishings, resulted in an increase of about 1.0 per cent over the month. Excise taxes levied by the Revenue Act of 1941 caused the further increase of about 0.2 per cent."

changes in consumption would have had as much effect on the movement of this index as on that of the United States index.

One change in policy was announced with the publication of the July index regarding direct taxes. Formerly, these taxes had been included in the index but in accordance with the Order-in-Council P. C. 6210, the prices of cigarettes and tobacco used in computing the official index do not include taxes imposed on June 24, 1942, under the Special War Revenue Act. An index comparable with the former index, including the tax, is still computed, together with the new index.

The Ministry of Labour which compiles the official cost of living index for Great Britain, describes this index[2] as representing changes in the cost of the average British workman's plane of living in 1914.

The International Labour Office makes the following comment on this procedure, in computing the British food cost index in wartime.[3]

> It is obvious, therefore, that apart from whatever changes occurred between 1914 and the present war in the composite commodity consumed by the average working-class family, the (British) index is unreliable as an index of the cost of food to the average family in war time. Rationing has considerably altered the kinds and amounts of goods bought by the people. Goods are no longer available in the same quantities as before the war, and the cost of living in war time is the cost of buying a different collection of goods from that bought previously.

The index has been criticized from two points of view: (1) on the ground that the increase in the average household expenditure on many of the items included in the index could not have been so large as the rise in the index, since rationing has reduced the amount of food and clothing available; (2) on the ground that the index excludes a number of miscellaneous foods which have risen in price more than those included in the index. These two criticisms, both of which have some validity, have been regarded by the British government as mutually cancelling out.

No information at present is available to the writers on the procedures used in introducing new types of textiles, clothing, and housefurnishings into the British cost of living index. It is known that 95 per cent of British clothing is now produced according to standard specifications in "utility models," and the index now inevitably follows the cost of some of these models.

[2] The weights are based on studies made in 1904 and 1908. The Ministry of Labour made a study of the expenditures of wage earners and lower salaried workers in 1937–38 for the purpose of revising these weights, but had not completed the revision by the time England declared war on Germany and the supply of consumers' goods in England as a result of the attack on Poland was necessarily greatly reduced.

[3] *Food Control in Great Britain*, International Labor Office, Montreal, 1942.

A number of adjustments have been made in the cost of living index for Sweden computed by the Social Board,[4] to allow for wartime changes in consumption. Thus, the weights have been reduced by ⅓ for cheese, ½ for eggs, ⅙ for flour, bread and cereals, and ¼ for sugar. While pork consumption has been reduced by not less than ⅔, the total meat ration approaches pre-war consumption. The amount of coffee allotted to households has been estimated to fall below usual consumption by only ¼ in view of the increased use of coffee substitutes. A corresponding reduction has been made in the weight for coffee in the index.

TABLE III

PER CENT INCREASE IN LIVING COSTS BY GROUPS OF ITEMS FROM AUGUST, 1939, TO AUGUST, 1942, IN CANADA, GREAT BRITAIN AND THE UNITED STATES

	All Items	Food	Clothing	Housefurnishings	Rent	Fuel and light	Misc.
Canada	16.8	30.5	20.0	16.7*	7.2	13.0	5.7
Great Britain	29.7	16.8	95.2	†	1.2	33.3	40.7
United States	19.1	34.0	24.7	21.1	3.2	8.9	10.7

* Includes services.
† Included with miscellaneous.

Many factors have been responsible for the disparity in the living cost curves in Great Britain, Canada, and the United States. Differences in the time and degree of active participation in the war, the price control policies adopted, and differences in self-sufficiency are inevitably reflected in the cost of living indexes of the three countries. The situation in Britain was most acute and costs in every major category of family expenditure showed far greater increases than occurred in either of the other two countries until the early part of 1942. At the beginning of this year, the British system of subsidizing production of consumers goods became an effective control and, coupled with lend-lease shipments, was instrumental in bringing about a decline in food costs. Food prices in the United States, however, and those in Canada which are to a large degree dependent on those in our country, were at a higher level in July, 1942, compared with the pre-war period, than were those in Great Britain. Costs for the other major groups continued at a relatively higher level than those in America and Canada. Table III shows the rise in each of the three countries by groups from August, 1939, to August, 1942.

[4] *Socialia Meddelanden*, November 10, 1941.

SELECTIVE SERVICE'S MEDICAL STATISTICS PROGRAM

By Oliver Harold Folk
Captain, Infantry, U. S. Army
Selective Service System

The selective service system operated for more than a year before the declaration of war, the first American attempt at peacetime conscription. During this period the primary function of the System was to procure a relatively small number of registrants for a peacetime training program. At that time, summary reports prepared by local boards on classification status and related information seemed to meet the necessities of the System. However, as the international situation developed into a global conflict, the necessity for more detailed and extensive information became apparent.

Registrants not deferred for occupational, dependency, or reasons other than physical or mental, are given a physical examination and, from the beginning of the Selective Service, National Headquarters has received copies of the reports of these physical examinations. These records had not been analyzed by the fall of 1941, at which time a Division of Research and Statistics was organized in National Headquarters and was charged with the responsibility of setting up a centralized system of individual records for all registrants. This system required the establishment of a procedure for receiving certain information relative to each registrant similar to the procedure for receiving a copy of the report of physical examination. This was accomplished so that forms for reporting statistical information could, in general, be made as an additional carbon copy of forms prepared for administrative purposes.

The medical statistics program of the Selective Service System is an integrated part of the entire statistical program, which is based on four essential forms: (1) a list of registrants from each local board showing the name, address, order and serial numbers, race and date of birth; (2) a report of the new and previous classification as soon as each registrant is classified or reclassified; (3) the occupational questionnaire, which shows citizenship, marital status, number of dependents, education, work status, employment class and specialized skills in selected occupations; and (4) the report of physical examination. The list of registrants and the occupational questionnaire are received for all registrants, ages 18 to 65. The classification report is received for only those registrants who are militarily liable, ages 20 to 45, and the

report of physical examination is received for only those registrants militarily liable who are not deferred for occupational, dependency, or reasons other than physical or mental.

Punch cards are prepared for each of the reports and contain an identification number for the registrant, which serves as a basis for matching several records for any one registrant. This identification number consists of the local board number plus the registrant's order number. The local board number is made up of the codes for the corps area, state, county, and local board within the county.

The examining procedure has been altered once and the physical standards have been lowered twice since the beginning of the Selective Service System. Initially, registrants not deferred for occupational, dependency, and other reasons were examined by local boards to determine their acceptability for military service. Those classed as available for general military service were sent to Army induction stations for a final-type physical examination which determined whether or not they were to be inducted for military service. About 28,000 examining physicians and some 10,000 examining dentists assisted the 6,443 local boards in performing the local board physical examination. The physical standards for the examination of Selective Service registrants were used by both local boards and Army induction stations. With the outbreak of war, the need for medical and dental services in the armed forces increased considerably, and on January 1, 1942, local boards were instructed to examine registrants for the purpose of eliminating only those with permanently disqualifying defects that are obvious to the examining physician and do not require instruments for their detection. Those not disqualified by the local board are sent to the induction station. On February 13, 1942, the minimum standards for teeth and eyes were lowered, and effective August 1, 1942, certain types of limited service men, Class 1-B, were acceptable.

Until the change in examination procedure on January 1, 1942, the local board prepared a physical examination report and the Army induction station prepared a separate report for each registrant examined. The analysis of these records, therefore, required the matching of both reports to get the final determination as to physical fitness of the registrant. The new procedure instituted a form in which the findings of the induction station, as well as the local board, were entered on one report. The reports used under the old examining procedure contained a limited amount of information in addition to the medical findings, which consisted of race, occupation, birthplace, birthdate, urban or rural community, medical history, and Selective

Service classification. The new reports contain, in addition to these items, education, years experience at present job, income, employment class, marital status, citizenship, and previous military service.

From the report in current use, much information is available which is needed for administrative uses. The analysis of these reports is divided into two parts: first, a progress reporting procedure whereby the registrant's identification, occupation, marital status, number of dependents, birthdate, race, medical defects, induction station, and examination date are punched currently as received, and summaries made monthly for operational and control purposes by national and state headquarters; and second, intensive studies of samples of these records which will provide information for planning and policy purposes. Current reporting is limited to forms used in the new procedure and samples for intensive study from this group of records will be limited to less than five per cent. However, the tabulation of each report received for registrants examined before January 1, 1942, will not likely be undertaken, and accordingly, a ten per cent sample has been drawn for detailed analysis.

In connection with the preparation of the procedure and instructions for the analysis of the medical reports, a survey of the existing medical nomenclatures currently in use revealed that most of them have been prepared for some specialized purpose, usually morbidity or mortality studies. Selective Service, however, examines a cross-section of the male population between 20 and 45 years of age, an age group which has a relatively low morbidity and mortality rate and, consequently, most of the diseases, defects and impairments found in these men are the hidden abnormal health conditions that are usually not included in morbidity and mortality codes. Also a classification of diseases, defects, and impairments for coding the results of the medical examination of Selective Service registrants must reflect the physical standards by which registrants are examined and the administrative needs of the System.

The medical nomenclature prepared by the Selective Service System is based primarily on the frequency and types of diseases, defects and impairments found at the time of examination of registrants. The specific diagnoses made at the time of examination of 16,000 registrants selected proportionately from each state, at random, were transcribed on to cards, which were then sorted into groups to bring together the defects of each large anatomical section or system of the body. The cards in each of the 27 groups were subdivided into smaller classifications. These detailed classifications were reviewed to eliminate over-

lapping, to clarify titles, to group the ill-defined, and to add new classifications where necessary. Classification titles were grouped and code numbers were assigned so that the code would be flexible enough to care for additional age groups and females, if necessary.

This code of 522 rubrics is used in connection with samples drawn for intensive and detailed studies for planning and policy purposes. An abridged code of 98 classifications is used for complete tabulation and the monthly reports. The 522 classifications are convertible to the 98 classifications. The medical code was tried out on a test sample of 30,000 reports and the results of the test were discussed with specialists for the purpose of remedying such of the defects in the code as were evident.

In the mass processing of these records, an inseparable part of the medical code is the index. The more than 38,000 local board examining physicians and dentists described the defects with considerable diversity of terminology. For consistency in reducing these diagnoses and, in many cases, symptoms, to code numbers, it was necessary to include in the index synonymous terms or expressions actually found on the reports of physical examination. This index consists of 15,000 terms and furnishes a guide for coding more than 95 per cent of the entries being received. Additional terms are being added to this index periodically.

The number of times each registrant has been examined is indicated on the medical record so that the most recent examination report for a registrant can be brought together with previous reports to eliminate duplication. The examining physician, after entering the detailed findings of the examination on the report, also entered a summary of the defects in the order of their seriousness for military service. Each defect recorded is tabulated, with the major defect so designated that it can be tabulated separately. With few exceptions, only one defect is coded for each of the 27 anatomical sections or systems of the body. Further, the coordination operation of defect coding and the tabulation of association tables of defects eliminate coding the cause in one section and the defect in another, or the symptom in one and the diagnosis in another. In addition to the reports received for registrants given a complete physical examination, reports have been received for registrants rejected because of the lack of educational qualifications and for those rejected for obvious physical defects, such as idiocy, epilepsy, feeble-mindedness, blindness, deafness, loss of arms or legs, or confinement in institutions for the insane.

Codes for non-medical items on the punch card have been arranged, wherever possible and feasible, to make these data comparable with existing data of a related nature. The occupational code is based on the "Convertibility List of Occupations and Industry Classifications" prepared in 1940 by the Division of Statistical Standards of the Bureau of the Budget and is convertible to occupational information of the Bureau of the Census and of the United States Employment Service.

While these medical records are being processed and analyzed for operational, policy, and planning purposes, this unique set of data presents a rare opportunity to persons interested in health conditions. Accordingly, from time to time, summaries of these medical reports will be released.

THE MEASUREMENT OF CAPACITY UTILIZATION

By Irving H. Siegel
U. S. Bureau of Labor Statistics

Information on the extent to which existing plant facilities are currently being utilized is indispensable to officials guiding our war production effort and to the many statisticians and economists engaged in the "battle of prediction." The usefulness of much of the available data, however, is qualified by the absence of adequate concepts of capacity and by other fundamental weaknesses which become more evident as maximum output is approached. No doubt, statistics compiled by or for the war agencies will supplement the data available from other sources, but they will surely not be accessible to all who could make use of them. This paper is devoted in part to a discussion of certain theoretical aspects of measurement which are of interest, not only to those compiling new series for capacity and utilization, but also to those who must use what they find. From this theoretical vantage ground, some published indications of capacity utilization are then subjected to examination.

Capacity, as defined in available short-term measures of utilization, may range from some achievable norm to a barely approachable ideal, from a hypothetical output under certain specified conditions to some subjectively determined optimum or maximum.[1] The concept is generally applied to the output of a plant as a whole; the measure of utilization is usually based on the final product of a plant or on the activity of a single department which is supposed to typify the activity of the entire plant. The estimate of capacity is generally fixed for a relatively long period, during which no allowance is made for changes in the volume of equipment or for changes in production methods, operating conditions, and the composition of output.

An important consequence of the methods of measurement is the possibility of attaining a level of capacity utilization in excess of 100 per cent. In such an event, the meaning of an indicator is obscure since the percentage corresponding to true "full" capacity is not given. It would seem to be a desirable alternative to define capacity as a flexible maximum which assumes either actual or other specified conditions not conducing to less than actual production. Such a definition yields a significantly calibrated scale, since 100 would always be the greatest

[1] For descriptions of various measures of capacity utilization, see, for example: E. G. Nourse and Associates, *America's Capacity to Produce* (1934); W. L. Thorp, "The Problem of Overcapacity," *Economic Essays in Honor of Wesley Clair Mitchell* (1935), pp. 479–495; and the "1940 Supplement" to the *Survey of Current Business*.

possible utilization percentage. The capacity *estimate* would continually be adjusted either for comparability with the actual production conditions in each period or for conformity to new ideal conditions. Such adjustments might be made to reflect changes in amount of plant and equipment; diversion or conversion of equipment from the manufacture of one item to another; changes in yield resulting, say, from improvements in raw materials and production techniques; changes in the character of output; changes in the number of shifts and length of workweek; and changes in the volume of work subcontracted. Clearly, many concepts of capacity, all exceeding actual production, are possible. The meaning of an indicator of capacity utilization thus depends on the explicit definition of capacity selected.

In a certain sense, the development of satisfactory short-term concepts of capacity and capacity utilization depends on the clarification of the *short-term* concept of production. The conventional measure of production in terms of units of finished product may afford but a crude indication of activity in a short period, particularly if the cycle of fabrication is longer than that period and if the flow of the product through successive manufacturing stages is interrupted by storage or for other reasons is not uniform. If appropriate data were available, it would seem preferable to measure output, not in terms of finished products, but in terms of subproducts of sufficiently short stages in the manufacturing process. The latter method would result in a closer correspondence between activity and output. The number of units of finished product would be regarded, in such a case, as the specific output of the "last" stage of fabrication instead of as the output of the entire plant.

The above concept of production has its analogue in capacity. For short periods, it may be more satisfactory to consider capacity utilization in various departments than in the plant as a whole. A measure of capacity in terms of the number of units of finished product would then be applied, not to the plant as a whole, but to the "last" department. Departmental measures of capacity utilization would be particularly useful in industries where it is desirable, for reasons already given, to measure output in terms of subproducts; where the character of productive activity varies seasonally; where the degree of integration of establishments in an industry is not uniform; and where varying amounts of subproducts are obtained from sources outside an industry. A measure of capacity utilization for an entire plant (or for an industry with establishments of dissimilar structure) could theoretically be derived from departmental production and capacity figures.

For practical purposes, it may be sufficient to measure departmental rates of operation in terms of man-hours, employment, or (similar) machine-hours and to measure plant or industry capacity utilization in terms of man-hours or employment. Thus, one might relate total man-hours actually worked to the maximum that could be worked in the same plant (i.e., in the same work space on the same installed equipment) under actual conditions (e.g., during the same number of shifts of comparable duration, etc.) or under superior conditions (e.g., in the same plant operated on a three-shift basis with experienced workers, etc.). Such measures correspond closely to activity and have the advantage or disadvantage of not being influenced by variations in productivity at different levels of operation. They would be particularly useful for war industries, which subcontract much of their work, have a diversified output, and may have insufficient experience to estimate capacity in physical quantities. Caution must be exercised, however, in the construction and interpretation of indicators based on labor[1] or machine-hours.

In the light of the foregoing discussion, some contemporary indications of capacity utilization will now be considered. No measures are currently available for war industries proper or for subsidiary industries manufacturing complex products (e.g., ordnance or machine tools). Because of the length of the production process and the prevalence of subcontracting, it would appear desirable in such industries to measure utilization departmentally or on a man-hour basis. But most references to operation rates—found, for example, in glowing accounts of local successes in the battle of production—involve finished goods. The concepts of capacity and production are generally loose. An increase in finished products accomplished through subcontracting is treated like an increase in production within the industry; and operation at merely 100 per cent of capacity would seem to be a disappointing performance.

The nebulous state of information on capacity utilization in war industries is evident from a few quotations from recent reports. One observer tells us, for example, that a certain company "estimates that it is turning out machines at the rate of five times its estimated capacity—but it isn't sure." Why? Because, the company says, "no one knows what his capacity is until he is doing everything possible with all the

[1] In view of the current labor shortages, for example, man-hours of different kinds and grades applied to different tasks seem less homogeneous than ever. It is important to distinguish between capacity man-hours for an activity pattern corresponding to actual production and capacity man-hours for an ideal pattern. If labor capacity were being estimated for three-shift operation, it would be a mistake to assume three times the capacity of the first shift. (See "Utilization of Plant Facilities Under National Defense Program," *Monthly Labor Review*, November 1941, pp. 1140–1147, and "Working Hours in War Production Plants, February 1942," *Monthly Labor Review*, May 1942, pp. 1061–1066.)

means at his disposal."[3] The same reporter cites a government armor plant which is now producing "far beyond its rated capacity."[4] In May, newspapers reported "capacity output steadily topped" in Canada's war plants. These plants, according to a digest of the Munitions Minister's remarks to Commons, were producing "above maximum rate." Furthermore, "explosive plants were producing as much as 75 per cent in excess of estimate and one gun plant was turning out two and a half times as many guns as it was thought it could produce."[5] The general manager of a plane company recently confided to newspapermen that "production is only 110 per cent of capacity: 'You see,' he added grinning, 'there is a WPB regulation that says we can't make more than 110 per cent of our allotted monthly quota of planes. If I were to tell you that we are producing more than 110 per cent, the WPB—while I don't guess it really would put me in jail—might not like it.'"[6]

Of the industries for which statistical series are available, steel is by far the most important. In war or peace, the fluctuations in the series for this industry are closely chronicled everywhere. The American Iron and Steel Institute measures of capacity utilization will, therefore, be considered in some detail. Very properly, the series, which are available weekly and monthly, are presented for separate "departments"—i.e., for major types of ingots and rolled products.

Since late in 1940, the measure of capacity utilization for all steel ingots has hovered near the 100 per cent level. Almost every rise or fall is deemed worthy of explanation, and it seems to be assumed that 100 is the true limit. But 100 is not a maximum, as is clear from the methods of measurement, from the achievements of individual companies, and from the records of the component varieties of steel. Consequently, even with current facilities, the industry might be expected to surpass 100 and continue to rise to a truer maximum.

The possibility of exceeding 100 per cent is not only implicit in the definition of "rated capacity" but is also conceded by the steel industry. "Rated capacity" represents "maximum output obtainable in normal operation, as based on tonnages actually obtained"; in its derivation, allowance is made for normal periodic interruptions.[7] Recognizing the "conservative" nature of such an estimate, the industry regards as feasible an annual output of 102.5 per cent of capacity.[8] But, if physical

[3] Paul McCrea, "The Machines Behind the Guns," *Nation's Business*, February 1942, p. 19.
[4] *Ibid.*, p. 27.
[5] *New York Times*, May 15, 1942.
[6] *New York Times*, May 19, 1942.
[7] *Steel Facts*, March 1940, p. 5.
[8] *Steel Facts*, February 1941, p. 3; May 1941, p. 2.

facilities alone were considered, this figure would seem to be an underestimate since individual companies have done much better for rather long periods. For example, a Midwest steel company exceeded its "theoretical capacity" by 19 per cent in 1941.[9] During the first four months of 1941, two-thirds of the 35 reporting companies equaled or exceeded their capacity estimates. Some of the plants also operated at comparable levels during many weeks of 1940. One company's rate was 105 per cent for 28 weeks of 1940; "and then, despite an increase in rated capacity at the year-end, continued to operate at an average of 190 per cent of capacity during 14 of the first 15 weeks of 1941." The 17 plants exceeding capacity surpassed it by 3.3 per cent.[10]

It is a curious fact that, despite the failure of the rate for all steel to exceed 100 per cent, the rates for two of the three major types have surpassed that level several times.[11] For example, the percentage for open-hearth steel, the most important type distinguished by the American Iron and Steel Institute, was 101.9 in March 1941, 100.0 in April, and 100.4 in May. The record of electric steel is more phenomenal. Its rate exceeded 100 in six months of 1941 and will approach, if not duplicate, this feat in 1942. In May 1941, when the rate for all steel was only 98.5, the rates for open-hearth and electric steel were 100.4 and 108.4, respectively. In June 1941, when electric steel reached a peak of 110.8, the average for all steel was only 98.1.

The failure of steel production to be maintained at full rated capacity or higher is too often attributed to physical factors entirely, like the desultory flow of scrap or the shutdown of a furnace for repairs. An important "deterrent" receiving too little attention is purely statistical—namely, the practice of revising capacity estimates at half-year intervals rather than as important changes occur. Recent periodic adjustments have tended to reduce the rates for open-hearth and electric steel and to increase the rate for Bessemer. The reduction, however, has outweighed the rise, forcing down the indicator for all steel in January and July. In reality, the figures for these months are more correct than the percentages for the rest of the year, which tend to be too high. Probably because of the infrequency of revision, however, there is a disposition to accept the inflated figures for all steel as proper indications of operating level. That the overstatement may be substantial is seen, for example, from the magnitude of the adjustment in July 1941; the open-hearth rate declined from 99.4 in June to 94.4 in July, the

[9] "Steel for Victory," February 1942 (a one-page addendum to *Steel Facts*).

[10] *Steel Facts*, May 1941, p. 2.

[11] The discussion in this paragraph and the following one is based on monthly releases of the American Iron and Steel Institute.

electric-steel rate declined from 110.8 to 85.7, and the rate for all steel declined from 98.1 to 93.3. Similarly, the January 1942 revision helped reduce the December 1941 rate for open-hearth steel from 99.0 to 95.4, for electric steel from 101.2 to 96.3, and for all steel from 97.9 to 94.7.

It is often assumed that the percentage of capacity utilization in steelworks is indicative of the rate of operation in fabricating departments. The degree of correspondence, however, depends on a multitude of factors, such as demand, balance of capacity, number of shifts and length of workweek, and inventory policy. The dispersion of operation rates is very wide, as may be seen from statistics for March 1942,[12] when the rate for steel ingots was 98.2. In that month, the rates for various steel products were as follows: standard rails, 52.0; hot-rolled strip, 54.8; drawn wire, 90.7; heavy structural shapes, 91.3; tool-steel bars, 90.0; plates (sheared and universal), 139.5; "all other" (miscellaneous), 190.9; and blackplate, 238.1. It is clear that, in rolling mills as well as in steelworks, it is not unusual nowadays for production to exceed estimated capacity. An interesting illustration of the irrelevance of certain capacity estimates under current operating conditions is afforded by a wide strip mill in a Pittsburgh steel company. Though "theoretically rated" at 70,000 tons a month, this mill accounted for 104,000 tons in March in spite of the variety of products scheduled on it (plates, sheets, strip, and blackplate).[13]

Another important industry represented by a current indicator of utilization is cotton goods—or, more precisely, cotton spinning. This measure, compiled by the U. S. Bureau of the Census, is based on spindle hours.[14] The capacity estimates employed in this measure are now obsolete, for they are based on a 5-day, 80-hour week, whereas yarn mills now operate on more intensive schedules. Thus, it is not surprising to find this series so buoyant since the industry's conversion to war production. In every month of 1941, the rate was well over 100, and, in the first quarter of 1942, the average was about 135. The significance of such rates is impalpable so long as the percentage corresponding to a truer estimate of capacity under contemporary conditions is not known. From the rates themselves, we have no way of anticipating the limit of future changes. It should also be noted that, because of the current lack of balance between carding and spinning equipment, a measure of

[12] American Iron and Steel Institute statistics for steel products are not available for public use after March 1942.

[13] *New York Times*, April 13, 1942. (The quantity 104,000, misprinted in the reference edition, has been verified as correct.)

[14] For current figures, see *Survey of Current Business*. The method of computation is shown in the Census report, *Cotton Production and Distribution: Season of 1941* (Bulletin 178), p. 41.

spindle activity may not even be representative of the entire yarn mill.

By way of summary, four main points of this paper are now repeated. First, available short-term indicators of capacity utilization must be interpreted in the light of the stated or unstated assumptions involved in their construction. Second, it is desirable to extend the concept of capacity from the plant to the department, particularly in war industries where subcontracting is employed to increase the volume of finished output. Third, it may be convenient, especially where production cannot be measured readily on a short-term basis, where subcontracting is general, and where the degree of plant integration is not uniform, to define capacity and capacity utilization in terms of man-hours or employment. Finally, it is desirable, however capacity is defined, to contrive a measure of utilization which cannot exceed 100. Refinements such as these would not only improve our statistical artillery for the present war but would also give us better tools for handling the inevitable problem of "excess capacity" in the post-war period.

PRELIMINARY POPULATION ESTIMATES BASED ON RATION BOOK APPLICATIONS

By T. J. Woofter, Jr.
Director of Research, Federal Security Agency

In May 1942 the estimated population of the United States was about 134,400,000. Roughly 97 per cent[1] of this number filed applications for War Ration Book One (Sugar Ration). If, therefore, it is assumed that a 2.8 per cent undercount in applications was evenly distributed in all areas, then a reasonable approximation to a population count is arrived at by adding 2.8 per cent to the recorded registrations. Such estimates compared with the April 1940 Census yield an idea of the volume and direction of population change, the increases including both excess of births over deaths and migration.

General analysis of the results in the light of other data on changes in employment gives some assurance that the estimates of population obtained by increasing ration applications by 2.8 per cent for states are fairly valid for such large units, but are subject to considerable margin of error in smaller populations.

The principal elements of the population which, theoretically, were not included in the registration and which account for the undercount are:

(a) Members of the armed forces except the small number who were eating in their own domestic quarters;

(b) Institutional populations, such as jails, hospitals, asylums, etc.;

(c) Persons habitually eating at restaurants, although some of these probably did register.

Since these groups were probably larger than the undercount, an undetermined amount of overlap and duplication occurred.

The operation of these factors seems to have been fairly even in large areas. The percentage contribution to the armed forces does not fluctuate widely from area to area. Institutional population is slightly variable from state to state, but there is a wide fluctuation of numbers eating at restaurants.

Other factors which would particularly affect the comparisons of local areas with the 1940 Census are:

(1) Institutional populations are highly concentrated in local areas;

[1] Incomplete application data were used as of September 15, 1942, at which time national total registrations were incomplete by the omission of some large New York and Pennsylvania counties and a few scattered counties in other states. When final figures by states are made available by the Office of Price Administration, they will show a slight variation from this preliminary total.

(2) Temporary laborers would register as of their residence in May 1942 rather than their usual place of abode;

(3) There may in some instances have been lack of exact correspondence in local areas covered by registration and Census enumeration areas.

It is for these reasons that a good many local estimators who have a considerable a priori knowledge of population change have reported mixed results in the use of sugar registration figures as a basis for population estimates. Some of these local estimates have been considered satisfactory and some subject to critical doubts. It must, of course, be noted in this connection that local estimates may at times be as wide of the mark as estimates based upon such registrations.

With this warning as to method, the following observations are submitted as to population change in states and large areas, as estimated by the addition of 2.8 per cent to ration book applications.[2]

According to the estimates by states, 31 states and the District of Columbia increased and 17 decreased from April 1940 to May 1942. The sum of the increase was 3,000,000, whereas the excess of births over deaths in these states was roughly 2,100,000, indicating an approximate interstate migration of about one and a half million people.

When the states are divided into three groups with respect to their relation to the war effort, the results appear as in Table I.

Group 1. The group of 15 states ranking at the top in dollar value of defense contracts and the District of Columbia showed very marked increases, except in Kansas and New York.

In the case of Kansas concentrated industrial activity was offset by continued decrease in the large rural population.

New York entered the war period with the largest actual number and the highest percentage of the population unemployed, and it was able from this reservoir and its natural increase to supply the demands of war industry and export several hundred thousand migrants. In addition, retail and wholesale trade and the non-durable goods industries were not markedly stimulated by war, a fact which contributed to the spectacular loss of about 270,000 in New York City alone.

These states are grouped in the Northeastern Seaboard, around the Great Lakes and on the West Coast. The aggregate increase in the District of Columbia and the 13 states of this group which increased was 2,700,000, or about 4.2 per cent—offset partially as we have noted by decreases.

[2] It should be emphasized that this method adds the members of the armed forces back into their community of origin, yielding a figure as to citizens of the community rather than those in actual residence at the time. To estimate the change in actual residence, the number of members contributed to the armed forces would have to be deducted from the estimates as shown in this article.

TABLE 1

POPULATION ENUMERATED APRIL 1940, ESTIMATED MAY 1942, AND INCREASE
(By states)

	Enumerated April 1940	Estimated May 1942	Increase
Group 1. Maximum war activity			
California	6,907,387	7,366,154	458,767
Connecticut	1,709,242	1,817,579	108,337
Illinois	7,897,241	8,211,041	313,800
Indiana	3,427,796	3,582,415	154,619
Kansas	1,801,028	1,766,898	−34,130
Maryland	1,821,028	1,946,069	125,041
Massachusetts	4,316,721	4,390,698	73,977
Michigan	5,256,106	5,709,322	453,216
New York	13,479,142	(Not available)	
New Jersey	4,160,165	4,264,536	104,371
Ohio	6,907,612	7,123,006	215,394
Pennsylvania	9,900,180	(Not available)	
Texas	6,414,824	6,656,272	241,448
Virginia	2,677,773	2,871,025	193,252
Washington	1,736,191	1,799,087	62,896
District of Columbia	663,091	754,000*	90,909
Group 2. Medium war activity			
Alabama	2,832,961	2,984,275	151,314
Arkansas	1,949,387	2,033,377	83,990
Delaware	266,505	282,513	16,008
Florida	1,897,414	2,001,000	103,586
Georgia	3,123,723	3,164,313†	40,590
Louisiana	2,363,880	2,504,730	140,850
Maine	847,226	855,774	8,548
Mississippi	2,183,796	2,200,056	16,260
Missouri	3,784,664	3,847,680	63,016
Nevada	110,247	132,133	21,886
North Carolina	3,571,623	3,661,623†	90,000
Oregon	1,089,684	1,096,805	7,121
Rhode Island	713,346	740,728	27,382
South Carolina	1,899,804	1,958,501	58,697
Tennessee	2,915,841	3,012,092	96,251
Utah	550,310	570,410	20,100
West Virginia	1,901,974	1,916,035	13,061
Wisconsin	3,137,587	3,140,877	3,290
Group 3. Negligible war activity			
Arizona	499,261	477,643	− 21,618
Colorado	1,123,296	1,116,819†	− 6,477
Idaho	524,873	491,191	− 33,682
Iowa	2,538,268	2,517,171	− 21,097
Kentucky	2,845,627	2,837,144	− 8,483
Minnesota	2,792,300	2,754,822	− 37,478
Montana	559,456	539,366	− 20,090
Nebraska	1,315,834	1,280,407	− 35,427
New Hampshire	491,524	487,901	− 3,623
New Mexico	531,818	518,376	− 13,442
North Dakota	641,935	611,622	− 30,313
Oklahoma	2,336,434	2,235,835	−100,599
South Dakota	642,961	605,835	− 37,126
Vermont	359,231	352,181	− 7,050
Wyoming	250,742	239,000	− 11,742

* Applications increased 3 per cent.

† Preliminary.

Group 2. Eighteen states with smaller dollar values of industrial contracts, some of which had troop concentrations which attracted civilian population, and most of which have high rates of natural increases constitute the second group. These states showed some increases, though in most instances not so marked as those in Group 1. The aggregate increase of these states was 900,000, or about 2.6 per cent.

Group 3. Fifteen states with negligible war activity all decreased, the aggregate decrease being 400,000, or about 2.5 per cent.

A high degree of concentration is indicated by the examination of a few metropolitan areas known to have had the maximum industrial activity. (These estimates include whole counties, sometimes several counties surrounding metropolitan cities.)

Over 90 per cent of the increase in the 13 increasing states in Group 1 was in 27 metropolitan areas.

In Illinois, for instance, about 85 per cent of the increase was in the Chicago area. Los Angeles, San Francisco Bay, and San Diego included most of California's increase. More than the Massachusetts state total increase was accounted for in the Boston-Lowell-Lawrence area. More than the total Connecticut increase was in New Haven and Hartford Counties.

This high degree of concentration in a few areas indicates that practically all of the interstate movement and a much larger short-range movement has poured into these areas.

In the second group of states increases were not quite so concentrated, but a fairly large proportion of the migration was into a few industrial centers.

The outstanding conclusions from the foregoing analysis are that in general the large industrial states have become larger and that the trend toward metropolitan concentration has been even more accelerated by the war activity.

RECENT DEVELOPMENTS IN CORRELATION TECHNIQUE

By Paul S. Dwyer
University of Michigan

Every author who has computed correlations and regressions involving a number of variables is aware that extensive computational work is demanded. During recent years certain advances in the computational technique have been introduced which tend to make this calculational work but a fraction of what it was previously for many types of problems. Space forbids an exhaustive review of these advances in this paper though the bibliography at the end should be useful to the student who desires to explore some of them. It is the primary aim of this paper *to integrate a number of the most useful of these developments into a series of compact and easy techniques for the numerical solution of the classical problems of multiple and partial correlation and regression.*

The development of the modern calculating machine and the adaptation of successive addition to the computation of summed products with machines which are primarily adding machines, such as the Hollerith tabulator, have revolutionalized the computation of the Pearsonian correlation coefficient of the zero order during the past decade. However it is not this topic but rather the computation from the correlation matrix of the quantities used in studying multiple and partial correlation which is the subject of this paper. Each technique outlined builds upon the correlation matrix. Of course many of the techniques are applicable, with slight adjustment, to more general least squares and regression problems but various considerations indicate that an introductory paper of this size should be limited to "correlation" techniques.

A necessary condition for the improvement of calculational technique has been the development of the modern computing machine with automatic division, automatic multiplication and the possibility of evaluation of such expressions as $a-bc-de-fg$ and $(a-bc-de-fg)/h$ as a single operation without copying and resetting. But the chief saving in computational time and energy results not so much from the fact that machines can perform these operations, since we do not gain too much if we use these machines in carrying out techniques which are designed for less refined machines, but in building up new techniques which utilize the mechanical developments to the fullest extent and in finding previously unknown general relationships which can be used in simplifying the calculational work. The introduction of an appropriate

notation, which leads one to the discovery of previously unknown important relationships, for example, is apt to be of much greater value than a mechanical improvement in machines.

The advances which are stressed in this paper are

1. An appropriate notation for the general problem.
2. An abbreviation of the Doolittle solution of the normal equations using machines which simplifies the routine and cuts down the calculational work enormously.
3. The identification of terms in this abbreviated Doolittle solution as correlation constants.
4. Concise methods for calculating the inverse of a matrix.
5. The use of the incomplete inverse as a substitute for the back solution.
6. The identification of terms in the incomplete inverse solution with regression values.

It should be remarked that this paper is really the fifth of a series of papers on modern calculational methods and the basic understanding would be clarified by a reading of papers (28) (29) (35) (36) in which the author has shown the development and theory of topics 1, 2, 4 mentioned above. The reader is also referred to the booklet by Smith (37) which has an aim similar to that of the present paper but which does not use the abbreviated Doolittle solution. The abbreviated Doolittle solution simplifies the presentation, and its understanding, greatly since the illustrations are not cluttered up with a lot of unessential details. Smith refers to the use of the incomplete inverse solution as a substitute for the back solution as the "added column" method and this label seems to be as good as any. The careful student is also advised to consult Deming (14) and Ezekiel (32).

NOTATION

A standard notation for the classical multiple correlation and regression constants is used which involves primary and secondary subscripts with a dot separating them. The secondary subscripts indicate the auxiliary variables involved in the regression while the primary subscripts indicate the specific variables under consideration. Thus if r_{ij} is the correlation between x_i and x_j, $r_{ij.kl}$ is the partial correlation between x_i and x_j with the variables x_k and x_l partialled out and, less precisely, $r_{ij.}\ldots$ represents the partial correlation between x_i and x_j resulting from the elimination of all the other variables of the regression under consideration. The multiple correlation of x_i with the remaining variables is then represented by $r_{i.}\ldots$ The values $\beta_{ij.}\ldots$ are

the standard regression coefficients where x_i is the predicted variable. The standard deviation of the variable (sample) is represented by σ_i and the standard deviation of residuals (sample) by $\sigma_{i,\cdots}$

The main calculational work in estimating the functions of the parent consists in calculating the functions of the sample and it is this aspect of the problem which is stressed in this paper. Estimates in the population may then be obtained with the use of appropriate formulas.

The author has built up a notation for the solution of the Doolittle method which utilizes primary and secondary subscripts in a fashion similar to that of the notation of classical correlation and regression. A complete familiarity with this notation, which is explained under the section devoted to the abbreviated Doolittle solution, is necessary if one is to understand the ideas presented rather compactly in the tables and in the later sections.

THE ABBREVIATED DOOLITTLE SOLUTION

The simplification of the calculational technique is due more to the use of a compact, simple, and streamlined solution of the normal equations than to any other single factor. This solution results from eliminating unnecessary entries from the classical Doolittle solution (35) but it also results from the elimination of unnecessary entries from the rival method of single division (28) which Aitken has called the "method of pivotal condensation" (12). Being the logical sequel of improvements in each of these methods, having its entries directly related to determinantal theory (29), being relatively simple to learn and use, being very compact and requiring a minimum of space, and having its entries immediately interpretable as multiple correlation and regression constants, it is unquestionably the method which should be used by those seeking numerical solutions to the classical problems of multiple and partial correlation.

The normal equations are denoted by

$$\sum a_{ij}x_i = a_{n+1,j} \qquad j = 1, 2, \cdots, n$$

where, in the problem under consideration $a_{ij}=r_{ij}$ and the x_i are the standard regression coefficients. The matrix of coefficients is symmetric since $r_{ij}=r_{ji}$ so that entries below the diagonal need not be recorded in setting up the problem.

The equations for $n=3$ are given symbolically in the first three lines of Table I, General. The first column presents the coefficients of x_1, the second of x_2, the third of x_3, while the fourth presents the constant term. The fifth column is a check column and is obtained by adding the en-

tries in the row. The entries below the main diagonal are enclosed in parentheses.

TABLE I

ABBREVIATED DOOLITTLE SOLUTION

Row	General					Illustration				
1	a_{11}	a_{12}	a_{13}	a_{1y}	a_{1c}	1.000	.313	.280	.495	2.088
2	(a_{21})	a_{22}	a_{23}	a_{2y}	a_{2c}	(.313)	1.000	.652	.550	2.515
3	(a_{31})	(a_{32})	a_{33}	a_{3y}	a_{3c}	(.280)	(.652)	1.000	.803	2.735
4	a_{11}	a_{12}	a_{13}	a_{1y}	a_{1c}	1.000	.313	.280	.495	2.088
5	1	b_{12}	b_{13}	b_{1y}	b_{1c}	1.0000	.313	.280	.495	2.088
6		$a_{22\cdot1}$	$a_{23\cdot1}$	$a_{2y\cdot1}$	$a_{2c\cdot1}$		.902	.564	.395	1.861
7		1	$b_{23\cdot1}$	$b_{2y\cdot1}$	$b_{2c\cdot1}$		1.000	.625	.549	2.174
8			$a_{33\cdot12}$	$a_{3y\cdot12}$	$a_{3c\cdot12}$			.569	.355	.924
9			1	$b_{3y\cdot12}$	$b_{3c\cdot12}$			1.000	.624	1.624
10	$\beta_{y1\cdot23}$	$\beta_{y2\cdot13}$	$\beta_{y3\cdot12}$	$a_{yy\cdot123}$		.271	.159	.624	.262	
11	$\beta_{c1\cdot23}$	$\beta_{c2\cdot13}$	$\beta_{c3\cdot12}$			1.271	1.159	1.624		

The first step in the solution consists in writing the equations and the check column in the form indicated. This is done in the illustration of Table I. The correlations used here are those obtained by Carver (27) from a study of body measurements of 1,000 male students entering the University of Michigan. The variables are in order, height, shoulder girth, chest girth, and weight.

The next step consists in writing the values a_{i1} in row 4 so that this row is identical with row 1. These values are then divided by the leading coefficient a_{11} and the results recorded in row 5. The entries in the fifth row are then $b_{i1}=(a_{i1}/a_{11})$. In the correlation problem row 5 is identical with row 4 since $a_{11}=1$. In general the entry in the check column obtained by division is identical, aside from rounding errors, with the sum of the entries to the left.

The next step consists in computing the values

$$a_{i2\cdot1}=a_{i2}-a_{i1}b_{21}\quad(\text{or } a_{i2}-a_{21}b_{i1} \text{ if one prefers})$$

in recording these entries in row 6, and in computing the values $b_{i2\cdot1}=(a_{i2\cdot1}/a_{22\cdot1})$ by dividing the entries of row 5 by the leading coefficient $a_{22\cdot1}$. This process of computing $a_{i2\cdot1}$ is carried out easily if the first row is covered with a ruler or pencil and the first column is covered similarly so that from the top visible entry a_{i2} one subtracts the product of a_{i1} directly beneath and b_{21}. Thus in the illustration $a_{32\cdot1}=.652-(.280)(.313)=.564$. It is important that the machine be used entirely

in this operation since even a hand calculator is capable of performing an $a-bc$ operation without recording and resetting the intermediate figures. The value $a_{s2\cdot 1}$ is used as a check. These entries are then divided by $a_{22\cdot 1}$ to get the elements of row 7. Reciprocals or locked divisors may be used depending on the calculating machine available. Also the values of the b's may be found by dividing the a's by $a_{22\cdot 1}$ as soon as each $a_{i2\cdot 1}$ is available, since machines are capable of performing $(a-bc)/d$ operations to the required number of significant places without resetting. As soon as the b's are determined, the check entry is computed and the sum checked.

The next step is similar. The entries of row 8 are obtained by the use of the formula

$$a_{i3\cdot 12} = a_{i3} - a_{i1}b_{31} - a_{i2\cdot 1}b_{32\cdot 1}$$

and the values of the entries in row 9 by the formula

$$b_{i3\cdot 12} = \frac{a_{i3\cdot 12}}{a_{33\cdot 12}}$$

or from a single operation by

$$b_{i3\cdot 12} = \frac{a_{i3} - a_{i1}b_{31} - a_{i2\cdot 1}b_{32\cdot 1}}{a_{33\cdot 12}}.$$

Calculation is facilitated if the first two rows and the first two columns are covered up. In the illustration, for example, one familiar with the meaning of the notation recognizes

$$a_{33\cdot 12} = 1.000 - (.280)(.280) - (.564)(.625) = .569$$

$$a_{43\cdot 12} = .803 - (.495)(.280) - (.495)(.625) = .355$$

$$b_{43\cdot 12} = \frac{.355}{.569} = \frac{.803 - (.495)(.280) - (.495)(.625)}{.569} = .624.$$

This process, continued in the general case until all the predicting variables are eliminated, constitutes an abbreviated form of the forward Doolittle solution. It is very easy to apply when thoroughly understood. Forms or templates may be constructed to assist in the selection of appropriate terms if desired.

An abbreviated form of the back solution is next in order and is presented in row 10. The value $x_3 = \beta_{43\cdot 12} =$ the value of $b_{43\cdot 12}$ in row 9 is placed in row 10 in the third column. Thus in the illustration the value of $x_3 (= \beta_{43\cdot 12})$ is .624. The value of x_2 is then found by solving the equa-

tion of row 7, $x_2+b_{23\cdot1}x_3=b_{42\cdot1}$ so that $x_2=b_{42\cdot1}-b_{23\cdot1}\beta_{43\cdot12}$. Now $\beta_{43\cdot12}$ is entered in row 10 directly under the value of $b_{23\cdot1}$ so that it is not difficult to find the terms and compute the value as a single operation on the machine. Thus in the illustration $\beta_{42\cdot13}=.549-(.625)(.024)=.150$. In a similar way it can be shown from row 5 that

$$\beta_{41\cdot23} = b_{41} - b_{31}\beta_{43\cdot12} - b_{21}\beta_{42\cdot13}$$

so that, in the illustration

$$\beta_{41\cdot23} = .405 - (.280)(.024) - (.313)(.150) = .271.$$

The abbreviated back solution in the more general problem proceeds in exactly the same fashion and requires but a single row.

The back solution can be checked by carrying the check column through the back solution in a similar manner. The entries are shown in row 11 and each is respectively one greater than the corresponding regression coefficient. The real check, however, consists in showing that the calculated coefficients satisfy the original normal equations. The form of the solution is ideally adapted to this type of check since it is easy to pick out the entries and compute

$$(1.000)(.271) + (.313)(.150) + (.280)(.024) = .405, \text{ etc.}$$

Actually some of us feel that the row sum check solution is unnecessary with modern abbreviated methods. In the interest of simplicity it is omitted from the illustrations of later sections.

More detail on the abbreviated Doolittle method cannot be given here since its merits are discussed elsewhere but enough must be given so that the reader understands the notation inasmuch as the contents of later sections are expressed largely in terms of this notation.

It should be noted that a variation of the abbreviated Doolittle method known as the Crout method is being used (30) by some. This method uses columns for the a's, rows for the b's, and so permits a compact solution. The mathematical methods of calculation and the formulas are identical with the methods outlined above but the entries are in different places so there is a somewhat different mechanical technique. An outline of the method, with variations in notation, addition of checks and double diagonal entries introduced by the present author, is presented in Table II.

The solution of the normal equations gives the regression coefficients. One may then use the formula

$$r_{4\cdot123} = \sqrt{\beta_{41\cdot23}r_{41} + \beta_{42\cdot13}r_{42} + \beta_{43\cdot12}r_{43}}$$

TABLE II

SECOND FORM OF ABBREVIATED DOOLITTLE SOLUTION

Variation of Crout Method

General					Illustration				
a_{11} a_{12} a_{13}	a_{12} a_{22} a_{23}	a_{13} a_{23} a_{33}	a_{14} a_{24} a_{34}	a_{1s} a_{2s} a_{3s}	1.000 .313 .280	.313 1.000 .052	.280 .052 1.000	.405 .650 .803	2.088 2.015 2.735
a_{11}	b_{12}	b_{13}	b_{14}	b_{1s}	1.000	.313	.280	.405	2.088
a_{21}	$\frac{1.000}{a_{22\cdot1}}$	$b_{23\cdot1}$	$b_{24\cdot1}$	$b_{2s\cdot1}$	.313	$\frac{1.000}{.902}$	.025	.540	2.174
a_{31}	$a_{32\cdot1}$	$\frac{1.000}{a_{33\cdot12}}$	$b_{34\cdot12}$	$b_{3s\cdot12}$	.280	.504	$\frac{1.000}{.500}$	.024	1.024
a_{41}	$a_{42\cdot1}$	$a_{43\cdot12}$	$a_{44\cdot123}$		.405	.405	.355	.202	
a_{s1}	$a_{s2\cdot1}$	$a_{s3\cdot12}$			2.088	1.901	.021		
$\beta_{41\cdot23}$	$\beta_{42\cdot13}$	$\beta_{43\cdot12}$			.271	.150	.024		

in obtaining the multiple correlation coefficient. In the illustration above, for example

$$r_{4\cdot123} = \sqrt{(.271)(.405) + (.150)(.650) + (.024)(.803)} = .850.$$

However it is possible, as was pointed out by Horst (1), to obtain the multiple correlation coefficient from the forward solution itself without the necessity of a back solution. For the term $a_{44\cdot123}$ is itself the square of the multiple alienation coefficient as is seen from the relation between the Doolittle solution and determinants (20)

$$r_{4\cdot123} = \sqrt{1 - \frac{\Delta}{\Delta_{44}}} = \sqrt{1 - a_{44\cdot123}}.$$

The value $a_{44\cdot123}$ of the illustration of Table I, $1-(.495)^2-(.405)(.540)-(.355)(.024) = .202$ is inserted in row 10. It is at once possible to obtain the multiple correlation coefficient $r_{4\cdot123} = \sqrt{1-.202} = .850$ without a back solution. The multiple correlation coefficient is also obtainable from the formula

$$r_{4\cdot123} = \sqrt{a_{41}b_{41} + a_{42\cdot1}b_{42\cdot1} + a_{43\cdot12}b_{43\cdot12}}.$$

The standard deviation of residuals is obtainable from the formula

$$\sigma_{44\cdot123} = \sigma_4\sqrt{\frac{\Delta}{\Delta_{44}}} = \sigma_4\sqrt{a_{44\cdot\cdots}}$$

if σ_1, the standard deviation of x_1, is known.

SOLUTION OF RELATED EQUATIONS

One forward solution may be used with different back solutions to solve a large number of systems of regression equations. These regression systems which have a common forward solution may be said to be "related."

Kurtz (9) appears to be the first who saw the application of the Doolittle method to the solution of related equations though he did not use an abbreviated form of the Doolittle solution nor did he incorporate Horst's idea of getting multiple correlations from the forward solution.

The solution of related equations using the abbreviated forward and back Doolittle solution is shown in Table III. The illustration is based on the Carver data from 1,000 students with x_1 = height, x_2 = shoulder girth, x_3 = chest girth, x_4 = waist girth, x_5 = right thigh girth, and x_6 = weight.

The abbreviated forward Doolittle solution is presented in the first 15 rows. The back solutions for different regression systems are shown in rows 16–30. The back solution for the six variable problem in which weight is estimated from the other variables is shown in row 20. The regression equation resulting from the omission of right thigh girth is shown in row 19 and is obtained by covering up the entire column 5 and completing the abbreviated back solution as indicated above. Row 18 is obtained by also omitting waist girth, row 17 also shoulder girth, etc. The sequence of equations 16–20 shows how the standard regression coefficients change with the introduction of an additional predicting variable.

The entries under column 6 are the values of $a_{66}\ldots$, the square of the alienation coefficient. Rows 21–24 exhibit a corresponding set of regression equations when x_5 is predicted from x_1, x_2, x_3, x_4 and rows 25–27 when x_4 is predicted from x_1, x_2, x_3, etc. The predicted variable is enclosed in each case by vertical lines. The different multiple coefficients are shown in column 7 and the multiple correlation coefficients in column 8. Standard deviation of residuals in the population can then be estimated by appropriate formulas.

THE IDENTIFICATION FORMULAS

It appears that the terms of the abbreviated Doolittle solution may be identifiable in terms of correlation and regression constants and, indeed, this is precisely the case. The identification is secured through the

TABLE III

SOLUTION OF RELATED EQUATIONS

Row							r^2	r
1	1.000	.313	.280	.182	.100	.405		
2	—	1.000	.052	.551	.018	.050		
3	—	—	1.000	.747	.093	.803		
4	—	—	—	1.000	.774	.804		
5	—	—	—	—	1.000	.812		
6	1.000	.313	.280	.182	.100	.405		
7	1.000	.313	.280	.182	.106	.405		
8		.902	.501	.497	.603	.405		
9		1.000	.025	.551	.024	.510		
10			.603	.385	.205	.353		
11			1.000	.077	.518	.024		
12				.432	.234	.201		
13				1.000	.512	.405		
14					.311	.128		
15					1.000	.375		
16	.405					.755	.245	.405
17	.323	.649				.483	.517	.710
18	.271	.150	.624			.262	.738	.850
19	.283	.100	.309	.405		.168	.832	.912
20	.310	.013	.253	.262	.375	.120	.880	.938
21	.166				.072		.028	.160
22	−.029	.621			.021		.370	.610
23	−.073	.300	.518		.408		.532	.720
24	−.017	.231	.151	.512	.341		.659	.812
25	.182			.007			.033	.182
26	.010	.551		.013			.307	.551
27	−.018	.128	.677	.432			.508	.754
28	.280		.622				.078	.280
29	.081	.625	.569				.431	.657
30	.313	.902					.098	.313

use of formulas which express the constants in terms of entries of the abbreviated Doolittle solution. Thus such formulas as

$$r_{i\cdot\ldots} = \sqrt{1 - \frac{\Delta}{\Delta_{ii}}} = \sqrt{1 - a_{ii\cdot\ldots}}$$

$$\sigma_{i\cdot\ldots} = \sigma_i\sqrt{\frac{\Delta}{\Delta_{ii}}} = \sigma_i\sqrt{a_{ii\cdot\ldots}}$$

$$\beta_{ij\cdots} = \frac{a_{ij\cdots}}{a_{jj\cdots}} = b_{ij} \qquad \text{if} \quad i > j$$

$$\beta_{ji\cdots} = \frac{a_{ji\cdots}}{a_{ii\cdots}} = \frac{a_{ij}}{a_{ii}} \qquad \text{if} \quad j > i$$

$$r_{ij\cdots} = \sqrt{\beta_{ij\cdots}\beta_{ji}} = \frac{a_{ij}}{\sqrt{a_{ii}\cdots a_{jj\cdots}}}$$

$$= \sqrt{\frac{a_{ij}\cdots b_{ij}}{a_{ii}}} = \sqrt{1 - \frac{a_{ii\cdots j,\ldots}}{a_{ii\cdots j}}}$$

permit the immediate identification of various entries of the forward abbreviated Doolittle solution with correlation constants. This is particularly true if the solution is augmented to record the values of $a_{jj\cdot 1}$, $a_{jj\cdot 12}$, $a_{jj\cdot 123}$, etc. in an adjacent column. This involves no additional work except recording since these quantities are computed in the process of computing $a_{jj\cdots}$. The expanded solution is shown in Table IV.

TABLE IV

THE AUGMENTED FORWARD SOLUTION

1.000		.313		.280		.182		.166		.495	
—		1.000		.052		.654		.015		.550	
—		—		1.000		.747		.093		.807	
—		—		—		1.000		.774		.904	
—		—		—		·		1.000		.812	
1.000	1.000	.313	1.000	.280	1.000	.182	1.000	.166	1.000	.495	1.000
1.000		.313		.280		.182		.166		.495	
		.902	.902	.604	.922	.497	.967	.603	.972	.495	.755
		1.000		.025		.551		.624		.540	
				.560	.509	.385	.093	.295	.021	.355	.481
				1.000		.077		.518		.624	
						.432	.432	.231	.408	.201	.262
						1.000		.612		.465	
								.341	.341	.128	.108
								1.000		.375	
										.120	.120

The entries in the added columns are successively a_{ii}, $a_{ii\cdot 1}$, $a_{ii\cdot 12}$, etc. These additional entries are useful in computing the conjugate regression coefficients and the partial correlation coefficients as well as a whole set of related multiple correlation coefficients. For the entries of the added columns are themselves the squares of the multiple alienation

coefficients (identical with those of Table III). Also every $b_{ij}\ldots$ is itself a regression coefficient since $b_{ij}\ldots=\beta_{ij}\ldots$, while every $a_{ij}\ldots$ has its $b_{ij}\ldots$ directly underneath it and its $a_{ii}\ldots$ at its right. Hence every term in the above solution is applicable to one or more of the identification formulas above.

In the illustration consider $a_{65\cdot1234}=.128$ and its auxiliary values $b_{65\cdot1234}=.375$ and $a_{66\cdot1234}=.168$. We have at once by substitution in the identification formulas

$$r_{6\cdot1234} = \sqrt{1-.168} = .912$$

$$\sigma_{6\cdot1234} = \sigma_6\sqrt{.168} = 7.0 \text{ lbs.} \quad \text{with} \quad \sigma_6 = 17.1 \text{ lbs.}$$

$$\beta_{65\cdot1234} = b_{65\cdot1234} = .375$$

$$\beta_{56\cdot1234} = \frac{a_{65\cdot1234}}{a_{66\cdot1234}} = \frac{.128}{.168} = .767$$

and

$$r_{65\cdot1234} = \sqrt{(.375)(.767)} = \frac{.128}{\sqrt{(.341)(.168)}} = \sqrt{\frac{(.128)(.375)}{.168}}$$

$$= \sqrt{1-\frac{.120}{.168}} = .530.$$

Every other set of entries can be "identified" likewise.

It becomes apparent that the essential entries of this solution of the normal equations (the abbreviated Doolittle solution) are themselves immediately identifiable in terms of correlation constants and are not merely auxiliary steps in the computation of those constants. It is true that these constants are not always the ones desired, for example, one does not get all the regression coefficients of the back solution of Table III directly from the entries of Table IV, but it is important to recognize that (1) the entries of the solution are themselves identifiable correlation constants and (2) these can be used, in many problems, in computing desired sequences of constants with a minimum of time and effort.

PARTIAL CORRELATIONS WITH THE SAME PRIMARY SUBSCRIPTS

The utility of this identification technique is shown by the solution of partial correlation when 1, 2, 3, 4, etc., variables are successively eliminated. Such a solution is outlined in Table V where as an illustration the correlations between height and weight for the Carver data are shown with shoulder girth, chest girth, waist girth and right thigh girth successively eliminated. The results show that the zero order cor-

TABLE V

COMPUTATION OF A SERIES OF PARTIAL CORRELATIONS HAVING THE SAME PRIMARY SUBSCRIPTS

										Variable
1	r	r	r	r	r					
—	1	r	r	r	r					
—	—	1	r	r	r					x_1 =
—	—	—	1	r	r					x_2 =
—	—	—	—	1	r	a	a	a	r	with
a	a	a	a	a	a	a	a	a	r	
1	b	b	b	b	b					x_3 =
	a	a	a	a	a	a	a	a	r	
	1	b	b	b	b					x_4 =
		a	a	a	a	a	a	a	r	
		1	b	b	b					x_5 =
			a	a	a	a	a	a	r	
			1	b	b					x_6 =
				a	a	a	a	a	r	
				(r)	a					

Illustration from Carver data

										Variable
1.000	.652	.554	.615	.313	.650					
—	1.000	.747	.603	.280	.803					
—	—	1.000	.774	.182	.801					x_1 = height
—	—	—	1.000	.166	.812					x_2 = weight
—	—	—	—	1.000	.495	a	a	a	r	with
1.000	.652	.554	.615	.313	.650	1.000	.495	1.000	.495	
1.000	.652	.554	.615	.313	.650					x_3 = shoulder girth
	.570	.380	.292	.070	.370	.002	.292	.578	.401	
	1.000	.071	.508	.132	.059					x_4 = chest girth
		.434	.237	−.042	.100	.892	.242	.328	.447	
		1.000	.546	−.097	.438					x_5 = waist girth
			.344	−.042	.116	.858	.200	.213	.558	
			1.000	−.122	.337					x_6 = right thigh girth
				.883	.274	.883	.274	.203	.615	
				(.046)	.205					

relation of .5 reduces to .4 and then increases to .65. Partial correlation studies of this sort, almost prohibitive by classical computational methods, are now easily and quickly found by arranging the variables represented by the primary subscripts at the right of the form, carrying

the solution through the elimination of the variables indicated by the secondary subscripts and with the continued use of the formula

$$r_{ij\cdot\ldots} = \frac{a_{ij\cdot\ldots}}{\sqrt{a_{ii\cdot\ldots}a_{jj\cdot\ldots}}}.$$

THE INVERSE OF A MATRIX

The inverse of the matrix is of great value in studying the regression problems. The application to groups of equations in which the predicting variables are the same was introduced by Gauss (14) and has been developed by various authors (4) (12) (13). The inverse of the correlation matrix is of great value, too, in expressing the standard errors of the regression coefficients. During recent years there have been a number of interesting methods developed for finding the inverse of a square symmetric matrix (4) (12) (13) (15) (19) (22) (28) (34) (36) (37). The method emphasized here owes much to the contributions of Aitken (12) although other authors (15) (36) (37) have made contributions. A more formal presentation and justification of the method than can be presented is found in the author's paper on "The Evaluation of Linear Forms" (36).

The method is illustrated in Table VI where the correlation coefficients are those of Table I. The identity matrix is set up on the right and the abbreviated Doolittle solution carried through the elimination of the four variables. The entries of the inverse matrix indicated by C_{ij} are then found by adding the products of a and b terms above. Diagonal

TABLE VI

INVERSE OF A SQUARE SYMMETRIC MATRIX

1.000	.313	.280	.495	1.000	0	0	0
—	1.000	.052	.050	0	1.000	0	0
—	—	1.000	.603	0	0	1.000	0
—	—	—	1.000	0	0	0	1.000
1.000	.313	.280	.495	1.000	0	0	0
1.000	.313	.280	.495	1.000	0	0	0
	.902	.601	.435	−.313	1.000	0	0
	1.000	.025	.540	−.347	1.100	0	0
		.509	.355	−.084	−.025	1.000	0
		1.000	.624	−.147	−1.098	1.757	0
			.263	−.271	−.150	−.624	1.000
			1.000	−1.031	−.007	−2.382	3.817
				1.401	−.001	.498	−1.031
				—	1.892	−.710	−.007
				—	—	3.213	−2.382
				—	—	—	3.817

entries, $i=j$, are found by summing products of terms in the i column while non-diagonal entries, $i\neq j$, are found by multiplying an a from the i-th column and a b from the j-th column and adding. Thus

$$C_{11} = (1.000)(1.000) + (-.313)(-.347) + (-.084)(-.147) + (-.271)(-1.034) = 1.401$$

and

$$C_{12} = C_{21} = (1.000)(0) + (-.313)(1.100) + (-.084)(1.098) + (-.271)(-.007) = -.091 \text{ etc.}$$

Smith (37) has called this method the "Added Column" method.

The adjoint can be found by multiplying each element of the inverse by the value of the determinant which is, in the illustration above, $a_{11}a_{22\cdot1}a_{33\cdot12}a_{44\cdot123} = (1.000)(.902)(.560)(.262)$.

REGRESSION COEFFICIENTS AND THEIR STANDARD ERRORS

The standard errors of the regression coefficients can be computed from the diagonal term of the inverse matrix since

$$\sigma_{B_{ij}\ldots} = \sqrt{\frac{a_{ij}\ldots}{n-1-m}}\, C_{jj}$$

where n is the number of observations and m is the number of independent variables. It appears at once that the diagonal terms only of the inverse solution are necessary and not the complete inverse. These can be written in the same row with the correlation coefficients and the value of $a_{ii}\ldots$

The use of the incomplete inverse is illustrated in Table VII where the regression coefficients of Table VI and their standard errors are computed. The first four columns and the first ten rows are identical with Table I. The last three columns carry out the inverse solution with the values of C_{11}, C_{22}, C_{33} entered in row ten, the values of

$$\frac{a_{44\cdot123}}{1000-1-3}, \quad \frac{a_{44\cdot123}}{996} C_{11}, \quad \frac{a_{44\cdot123}}{996} C_{22}, \quad \frac{a_{44\cdot123}}{996} C_{33}$$

in row 11, and the standard errors in row 12.

The values of the regression coefficients themselves may be found with the use of the inverse without the necessity of a back solution by multiplying the b's of column 4 by the a's of column 5, 6, 7 in turn.

Thus

$$\beta_{41\cdot 23} = (.495)(1.000) + (.549)(-.313) + (.024)(-.084) = .271,$$
$$\beta_{42\cdot 13} = (.495)(0) + (.549)(1.000) + (.024)(-.025) = .150,$$
$$\beta_{43\cdot 12} = (.495)(0) + (.549)(0) + (.024)(1.000) = .024.$$

TABLE VII

REGRESSION COEFFICIENTS AND THEIR STANDARD ERRORS FROM THE *INCOMPLETE INVERSE*

x_1	x_2	x_3	x_4			
1.000	.313	.280	.495	1.000	0	0
—	1.000	.652	.650	0	1.000	0
—	—	1.000	.803	0	0	1.000
1.000	.313	.280	.495	1.000	0	0
1.000	.313	.280	.495	1.000	0	0
	.902	.564	.495	−.313	1.000	0
	1.000	.625	.549	−.347	1.109	0
		.569	.353	−.081	−.025	1.000
		1.000	.624	−.147	−.068	1.757
.271	.150	.024	.262	1.121	1.706	1.757
			.000263	.000295	.000172	.000102
				.017	.0022	.024

MORE IDENTIFICATION FORMULAS

More identification formulas are now available since, as Horst (2) and Waugh (8) have shown, it is possible to compute from the inverse matrix (*a*) all multiple correlations of order *n*-1, (*b*) all partial correlations of order *n*-2, (*c*) all regression coefficients of order *n*-2, (*d*) the standard errors of regression coefficients of order *n*-2 with the use of the identification formulas

$$r_{i\ldots} = \sqrt{1 - \frac{1}{C_{ii}}}$$

$$r_{ij\ldots} = \frac{-C_{ij}}{\sqrt{C_{ii}C_{jj}}}$$

$$\beta_{ij\ldots} = -\frac{C_{ij}}{C_{ii}}$$

$$\sigma_{\beta ij\ldots} = \sqrt{\frac{C_{ii}C_{jj} - C_{ij}^2}{n'C_{ii}}}$$

where n' is the number of degrees of freedom.

More than that it can be shown that the a entries on the right hand side of Table VI are themselves regression coefficients and more than that they are the coefficients needed for a complete regression equation. Thus the coefficients of row 12 of Table VI are the negatives of the coefficients of row 18 of Table I (and row 10 of Table VII) and are the negative coefficients of the $4\cdot123$ system. Similarly the coefficients of row 9 of Table VI are the negative coefficients of the $3\cdot12$ system, etc. It thus appears that the entries of the inverse solution are also identifiable as correlation constants.

SUCCESSIVE MULTIPLE CORRELATION CONSTANTS FROM THE INCOMPLETE INVERSE

We are now ready for the illustration which shows the real value of many of the advances outlined above. One can build up values of regression coefficients, their standard errors, multiple alienation and correlation coefficients, and the standard deviation of residuals for related groups of equations readily from the incomplete inverse solution.

The illustration used is that of Table III with the incomplete inverse added. The regression coefficients themselves are computed easily by forming the summed successive products of the b's of column 0 and the a's of incomplete inverse. Thus

$$\beta_{01} = .495(1.000) = .495$$
$$\beta_{01\cdot1} = .495 + (.549)(-.313) = .323$$
$$\beta_{01\cdot12} = .323 + (.024)(-.084) = .271 \quad \text{etc.}$$

This process is most easily carried out and it is only necessary to record each entry as the final term is built up. The values in column 0 are the values $a_{ij}\ldots$ of Table III. The values $C_{ij\cdot j}$ are the corresponding diagonal elements of the inverse matrices of order 1, 2, 3, 4, 5, respectively. They likewise are easily obtained from the products of "a" and "b" terms above them. Thus

$$C_{11\cdot1} = (1.000)(1.000) = 1.000$$
$$C_{11\cdot2} = C_{11\cdot1} + (-.313)(-.347) = 1.100 \quad \text{etc.}$$

From these values it is possible to calculate the standard errors of the regression coefficients on the left.

The values $r_0^2\ldots$, $r_0\ldots$, and $\sigma_{e_0}\ldots$ are given at the right. Corresponding values can be readily computed for all the related regressions of Table III.

TABLE VIII

SUCCESSIVE MULTIPLE CORRELATION CONSTANTS FROM THE INCOMPLETE INVERSE

Illustration from Carver data

1.000	.313	.250	.152	.106	.495	1	0	0	0	0	x_0 = weight		
—	1.000	.652	.554	.615	.650	0	1	0	0	0	n = 1000		
—	—	1.000	.747	.693	.803	0	0	1	0	0	M_0 = 139.3 lbs.		
—	—	—	1.000	.774	.804	0	0	0	1	0	σ_0 = 17.07 lbs.		
—	—	—	—	1.000	.812	0	0	0	0	1			
1.000	.313	.250	.152	.106	.495	1.000	0	0	0	0	x_1 = height		
1.000	.313	.250	.152	.106	.495	1.000	0	0	0	0			
	.902	.564	.497	.563	.495	−.313	1.000	0	0	0	x_2 = shoulder girth		
	1.000	.625	.551	.624	.549	−.347	1.109	0	0	0			
		.569	.385	.295	.353	−.054	−.625	1.000	0	0	x_3 = chest girth		
		1.000	.677	.518	.624	−.147	−1.098	1.757	0	0			
			.432	.234	.201	.017	−.128	−.677	1.000	0	x_4 = waist girth		
			1.000	.542	.465	.109	−.298	−1.567	2.315	0			
				.341	.128	.047	−.231	−.151	−.542	1.000	x_5 = right thigh girth		
				1.000	.375	.138	−.677	−.443	−1.559	2.933			
.495					.735	1.000					.245	.495	14.53 lbs.
.323	.549				.483	1.109	1.109				.517	.719	11.86 lbs.
.271	.159	.624			.262	1.121	1.795	1.757			.738	.859	8.74 lbs.
.293	.099	.309	.465		.168	1.126	1.833	2.818	2.315		.832	.912	7.00 lbs.
.310	.013	.253	.262	.375	.120	1.133	1.990	2.855	3.176	2.933	.880	.938	5.91 lbs.
β_{01}					a_{01}	$C_{11\cdot1}$					$r^2_{0\cdot1}$	$r_{0\cdot1}$	$\sigma_{e_{0\cdot1}}$
$\beta_{01\cdot2}$	$\beta_{02\cdot1}$				$a_{02\cdot1}$	$C_{11\cdot2}$	$C_{22\cdot2}$				$r^2_{0\cdot12}$	$r_{0\cdot12}$	$\sigma_{e_{0\cdot12}}$
$\beta_{01\cdot23}$	$\beta_{02\cdot13}$	$\beta_{03\cdot12}$			$a_{03\cdot12}$	$C_{11\cdot3}$	$C_{22\cdot3}$	$C_{33\cdot3}$			$r^2_{0\cdot123}$	$r_{0\cdot123}$	$\sigma_{e_{0\cdot123}}$
$\beta_{01\cdot234}$	$\beta_{02\cdot134}$	$\beta_{03\cdot124}$	$\beta_{04\cdot123}$		$a_{04\cdot123}$	$C_{11\cdot4}$	$C_{22\cdot4}$	$C_{33\cdot4}$	$C_{44\cdot4}$		$r^2_{0\cdot1234}$	$r_{0\cdot1234}$	$\sigma_{e_{0\cdot1234}}$
$\beta_{01\cdot2345}$	$\beta_{02\cdot1345}$	$\beta_{03\cdot1245}$	$\beta_{04\cdot1235}$	$\beta_{05\cdot1234}$	$a_{05\cdot1234}$	$C_{11\cdot5}$	$C_{22\cdot5}$	$C_{33\cdot5}$	$C_{44\cdot5}$	$C_{55\cdot5}$	$r^2_{0\cdot12345}$	$r_{0\cdot12345}$	$\sigma_{e_{0\cdot12345}}$

CONCLUSION

There are many additional points which should be discussed but this paper is already sufficiently long. In summing up this portion of the material, I submit the thesis that the entries of the abbreviated Doolittle solution are so useful in finding multiple and partial coefficients of correlation and regression that the standard computational formulas should be expressed in terms of these entries rather than in terms of determinants or other functions of the correlations.

REFERENCES

The bibliography given is that presented at the 103rd Annual Meeting of the American Statistical Association in joint session with the Institute of Mathematical Statistics, New York, December 28, 1941 (with slight corrections and additions) and covers a somewhat wider field than does the present paper.

1932

1. P. Horst. "A Short Method for Solving for a Coefficient of Multiple Correlation." *Annals of Math. Stat.* 3 (1932) pp. 40–44.
2. P. Horst. "A General Method of Evaluating Multiple Regression Constants." *Jour. of Am. Stat. Assn.* 27 (1932) pp. 270–278.
3. A. C. Aitken. "On the Evaluation of Determinants, the Formation of their Adjugates, . . ." *Proc. of Edinburgh Math. Soc.* Second Series, Vol. 3 (1932) pp. 207–219.

1934

4. R. A. Fisher. *Statistical Methods for Research Workers.* 5th Edition, 1934. Oliver and Boyd.
5. R. Frisch. *Statistical Confluence Analysis by Means of Complete Regression Systems.* Publication no. 5 from University Institute of Economics, Oslo, 1934.

1935

6. M. Roff. "Some Properties of the Communality in Multiple Factor Theory." *Psychometrika.* 1 (1936) pp. 1–6.
7. G. Strecker. "On Evaluating a Coefficient of Partial Correlation." *Annals of Math. Stat.* 6 (1935) pp. 143–145.
8. F. V. Waugh. "A Simplified Method of Determining Multiple Regression Constants." *Jour. of Am. Stat. Assn.* 30 (1935) pp. 694–700.

1936

9. A. K. Kurtz. "The Use of the Doolittle Method in Obtaining Related Multiple Correlation Coefficients." *Psychometrika.* 1 (1936) pp. 45–51.
10. H. D. Griffin. "A Further Simplification of the Multiple and Partial Correlation Process." *Psychometrika.* 1 (1936) pp. 219–228.
11. F. V. Waugh. "The Analysis of Regression in Subsets of Variables." *Jour. of Am. Stat. Assn.* 31 (1936) pp. 729–738.

1937

12. A. C. Aitken. "Studies in Practical Mathematics, I. The Evaluation with Applications of a Certain Triple Product Matrix." *Proc. Roy. Soc. Edinburgh.* 57 (1937) pp. 172–181.
13. P. S. Dwyer. "The Simultaneous Computation of Groups of Regression Equations and Associated Multiple Correlation Coefficients." *Annals of Math. Stat.* 8 (1937) pp. 224–231.

1938

14. W. E. Deming. "Some Notes on Least Squares." The Graduate School, *Department of Agriculture, Washington,* 1938.
15. L. R. Tucker. "A Method for Finding the Inverse of a Matrix." *Psychometrika.* 3 (1938) pp. 189–197.
16. F. L. Wren. "The Calculation of Partial and Multiple Coefficients of Regression and Correlation." *Jour. Ed. Psy.* 29 (1938) pp. 695–700.

1939

17. W. Ledermann. "On a Shortened Method of Estimation of Mental Factors by Regression." *Psychometrika.* 4 (1939) pp. 109–116.
18. P. S. Dwyer. "The Contribution of an Orthogonal Multiple Factor Solution to Multiple Correlation." *Psychometrika.* 4 (1939) pp. 163–171.
19. J. W. Dunlap. *Workbook in Statistical Method.* Prentice-Hall, New York, 1939.

1940

20. L. Guttman. "Multiple Rectilinear Prediction and the Resolution into Components." *Psychometrika.* 5 (1940) pp. 75–99.
21. O. Reiersøl. "A Method for Recurrent Computation of All the Principal Minors of a Determinant, and its Application to Confluence Analysis." *Annals of Math. Stat.* 11 (1940) pp. 193–198.
22. W. D. Baten. "A Solution of Normal Equations Giving the Standard Errors of the Constants." *Jour. of Ag. Research.* 61 (1940) pp. 237–240.
23. P. S. Dwyer. "The Evaluation of Multiple and Partial Correlation Coefficients from the Factorial Matrix." *Psychometrika.* 5 (1940) pp. 211–232.
24. M. Roff. "Linear Dependence in Multiple Correlation Work." *Psychometrika.* 5 (1940) pp. 295–298.
25. J. C. Flanagan. "A Successive Approximation Solution for Prediction Problems Involving a Large Number of Variables." *Proceedings of Educational Research Forum.* pp. 75–79. Published 1940 by International Business Machines Corporation.
26. W. H. Stead and others *Occupational Counseling Techniques.* American Book Co., 1940. Especially the Wherry-Doolittle Method, pp. 245–250.

1941

27. H. C. Carver. *Anthropometric Data.* Edwards Brothers, Ann Arbor, 1941.
28. P. S. Dwyer. "The Solution of Simultaneous Equations." *Psychometrika.* 6 (1941) pp. 101–129.
29. P. S. Dwyer. "The Evaluation of Determinants." *Psychometrika.* 6 (1941) pp. 191–204.

30. P. D. Crout. "A Short Method for Evaluating Determinants and Solving Systems of Linear Equations with Real or Complex Coefficients." American Institute of Electrical Engineers, 33 West 39 Street, New York City, or *Marchant Methods, MM-182.* Sept., 1941. Marchant Calculating Machine Co., Oakland, California.
31. P. G. Hoel. "On Methods of Solving Normal Equations." *Annals of Math. Stat.* 12 (1941) pp. 354–359.
32. M. Ezekiel. *Methods of Correlation Analysis.* Wiley (1941). Second Edition.
33. P. Horst (Editor). *The Prediction of Personal Adjustment.* Bulletin 48 (1941) of the Social Science Research Council. Especially the material by L. Guttman on Mathematical and Tabulation Techniques, pp. 251–364.
34. M. D. Bingham. "A New Method for Obtaining the Inverse Matrix." *Jour. of Am. Stat. Assn.* 36 (1941) pp. 530–534.
35. P. S. Dwyer. "The Doolittle Technique." *Annals of Math. Stat.* 12 (1941), pp. 449–458.
36. P. S. Dwyer. "The Evaluation of Linear Forms." *Psychometrika.* 6 (1941), pp. 355–365.
37. J. H. Smith. *Added Column Method for Computing Multiple Correlation Constants.* First Edition December 1941. University of Chicago Press.

USE OF THE DISCRIMINANT FUNCTION FOR MORE THAN TWO GROUPS

By Besse B. Day and Marion M. Sandomire
U. S. Department of Agriculture

Differentiating between two or more populations by means of a single measurement has now become a standardized process, with well-developed fundamental theory. Conclusions thus reached, however, are likely to be inadequate. Usually, two or more measurements yield much more information than does one about the populations being studied. When two or more measurements are available, therefore, the question arises, How may these be compounded to obtain the best possible function for discriminating the populations under study? The discriminant function proposed by Fisher (4) in 1936 and further outlined in the 7th edition of his *Statistical Methods* (6) affords a procedure for obtaining the best linear function for this purpose.

This technique has been applied in widely differing fields as craniometry (1), taxonomy (4), plant selection (12), soil science (2), material psychology (13), eye color (10), and finance (3). As Fisher pointed out (7), considerable work in theory remains to be done. An analysis-of-variance test for significant discrimination of two groups of samples has been solved (5) by an adaptation of Hotelling's T^2, thus determining the comparative effectiveness of two discriminant functions. Such an exact test for a case of three or more groups with small samples is not as yet available. However, Hsu (9) has developed such a test for large samples. In addition, methods for calculating the standard errors and the significance of the individual coefficients remain undetermined.

Measurement data on Whitetail deer were generously placed at our disposal, in advance of their publication, by Barry C. Park, wildlife specialist of the U. S. Forest Service. Over a period of years, data on deer killed by hunters on the Alleghany National Forest, in Pennsylvania, had been collected under Park's direction. The data were of two types: Age classes as determined by a dental formula, and direct measurements of the body and (in the case of male deer) the antlers. As stated by Park, the dental-formula method of estimating age class, while considered accurate, is impractical for general use because only a highly trained person can apply it, and prying the jaws apart after freezing is difficult. Hence there was need for a simpler method.

The discriminant function method would seem to be more appropriate than the regression method since age (the dependent variable) can be obtained in broad classes, only, by use of the dental formula. The

present paper is concerned with the application of the discriminant function to estimating age class of deer. We shall present some intuitive methods, without proof, for overcoming the statistical limitations mentioned.

DESCRIPTION OF DATA

In the original data for age of deer taken by the dental-formula method, in the period 1936–39, four age classes were distinguished: Fawns, about 9 months old; young, about 1½ years old; mature, 2½ to 4½ years old; and overmature, 5½ years old or older. The herd studied included few very old animals. The data taken on deer killed in 1936 and some in 1937 did not include body measurements other than weight, and therefore were not used in the study except as a check on those functions obtained from weight and antler measurements. The sizes of the samples used in calculating the discriminant functions are given in Table I, the four age classes being designated by roman numerals I to IV. In the same table are listed the several characters measured, seven for males and three for females, and their mean values by sex and age class, and the standard deviations within classes. There were no significant differences in the measurements from year to year.

TABLE I

NUMBER OF DEER AND MEANS OF DEER MEASUREMENTS, BY SEX AND AGE CLASS*

Sex and measurement	9 months I	1½ years II	2½–4½ years III	5½ years and over IV	Standard deviation within age classes
Male					
Number	*104*	*94*	*170*	*164*	.
x_1, Dressed weight (lb.)	54.3	95.4	118.5	135.0	13.4
x_2, Length of body (cm.)	140.9	167.8	179.9	187.3	9.6
x_3, Length of hind foot (cm.)	43.3	47.5	49.0	50.3	1.8
x_4, Spread of antler (cm.)	—	18.0	34.2	41.5	7.1
x_5, Circumference of main beam (cm.)	—	4.8	7.6	9.0	1.3
x_6, Length of antler (cm.)	—	15.9	31.8	39.6	6.8
x_7, Number of points	—	2.9	6.5	7.7	1.7
Female					
Number	*111*	*54*	*150*	*150*	...
x_1, Dressed weight (lb.)	51.2	60.0	90.0	98.8	9.2
x_2, Length of body (cm.)	138.3	157.1	161.3	168.0	8.9
x_3, Length of hind foot (cm.)	41.6	45.7	46.1	47.6	1.9

* Supplied by Park from the original study (11).

STATISTICAL PROCEDURE

Effectiveness of single character. Despite the fact that several characters vary materially with age, considerable error might result in estimating the age of an individual deer from any one of them. Only if the measurements were average would they afford in themselves a dependa-

ble index of age. Frequency distributions showed the considerable variation in measurements for the deer in any age group (11).

If the arrays of the several measurements by age class are homoscedastic, the procedure is straightforward for the limits of the age classes fall midway between the group means. Then division of half the difference of the means for two consecutive age classes by the standard deviation within groups yields what might be considered a normal deviate. Its corresponding probability, P, found in certain tables, is the probability of an observation falling outside the range $-x$ to $+x$, where x is the deviate. In the consideration of any two successive age groups, only one tail of the probability curve is involved. Then we take one-half of P for the chance of a deer's falling either above or below the midpoint between the averages for two successive age groups, that is, above the smaller or below the larger age class. For age classes II and III, the total probability of making a mistake is the sum of two probabilities. For example, the chance of making a mistake in estimating the age of a $1\frac{1}{2}$-year deer is the probability of classing it as less than $1\frac{1}{2}$ years plus that of classing it as over $1\frac{1}{2}$ years.

The calculated probabilities for each character measured appear in Table IV under "Single measurement." For both males and females, dressed weight proved to be a better basis for estimate than either of the two other body measurements, although still not very dependable. For males, if dressed weight alone were used, there would be 6 chances in 100 of putting a 9-month deer in the $1\frac{1}{2}$-year class or above, 27 chances of misclassifying a $1\frac{1}{2}$-year, 47 for a $2\frac{1}{2}$–$4\frac{1}{2}$-year, and 20 for an overmature animal. The three antler measurements other than number of points were found to be better indicators of age than any of the three body measurements.

Compounds of two or more characters. We wish then to find the expression which will best characterize the age of individual deer, that is, the compound of measurements which will show the maximum change with age group relative to the variance of the compound within the age groups. If a linear function of the several deer measurements which shows the most distinct variation with age be represented by the linear expression

$$\phi = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_p x_p$$

then the coefficients $\lambda_1, \lambda_2, \cdots, \lambda_p$ are the solutions of the equations

$$\begin{cases} \lambda_1[x_1^2] + \lambda_2[x_1x_2] + \cdots + \lambda_p[x_1x_p] = B_1 \\ \lambda_1[x_1x_2] + \lambda_2[x_2^2] + \cdots + \lambda_p[x_2x_p] = B_2 \\ \vdots \\ \lambda_1[x_1x_p] + \lambda_2[x_2x_p] + \cdots + \lambda_p[x_p^2] = B_p \end{cases}$$

where $[x_p x_{p'}]$ is a simplified scheme for writing the pooled sums of squares and of products within the age classes. The λ's may be thought of as the weights of the variables in the function. The sums of squares and cross products (5; 6, art. 49.2) without regard to group means might have been used instead, and the resulting coefficients would have been proportional to the λ's obtained here.

The B's, which constitute the right-hand side of the equations, are the regressions of the means of the measurements on age. Thus,

$$B_p = \frac{\sum_{A=1}^{IV} (\bar{x}_{pA} - \bar{x}_p)(a_A - \bar{a})}{\sum_{A=1}^{IV} (a_A - \bar{a})^2}$$

where $\bar{x}_{pA}$ is the mean of measurement p for age A. This is a generalization of the procedure for discriminating two groups by use of the difference between the two means. The average difference among groups with respect to any one measurement is thus represented by the slope of the function of age (4, p. 185).

When the slope is to be used as a measure of the average change, it is desirable that the function of age be linear. The mean values for each character (Table I) were plotted first with rectangular coordinates, consecutive age groups being equally spaced. To overcome the nonlinear character of the relationships, a logarithmic transformation of the age values 1, 2, 3, and 4 was introduced.

Following Barnard (1) in the preliminary selection of those characters which would yield the best discriminating function, "Student's" t was calculated for a comparison of measurement values for the extreme ages, II and IV, since II was the youngest age for which all measurements were available. The resulting values for male deer appear in the first line of Table II. Next, the effect of each measurement, in turn, was eliminated from each of the other measurements, and the t values were recalculated for these adjusted measurement values.

Since the association between characters is considerable, the selection was to be based on greatest independent variation. This would consist of a combination of those characters having the greatest influence on all others and those having the least association with the others. In the case of male animals, length of antler seemed to have the greatest influence on the other measurements and length of hind foot seemed to be the most independent (Table II). After the removal of the effect of these two, the values of t for each of the other adjusted measurements

were recalculated; this process was repeated after removal of the effect of dressed weight as well. The work required for calculating the t's corresponding to all combinations of two measurements removed was prohibitive at the time.

It may be noted, in the upper portion of Table II, that removal of the effect of each of the three body measurements affected the values for the other body measurements but left those for the antler measurements little changed; and that removal of the effect of each of the four antler measurements had a marked effect on the other antler values and also on the first two body-measurement values.

TABLE II

Part 1: t values for comparison of age classes II and IV for each male deer measurement with effects of indicated measurement removed

Measurement effect removed	Dressed weight	Length of body	Length of hind foot	Spread of antler	Circumference of main beam	Length of antler	Number of points
. .	18.35	12.49	9.03	22.88	20.40	21.40	22.72
Dressed weight	. .	2.22	4.30	16.80	14.10	17.02	18.24
Length of body	13.60	.	6.21	19.04	17.40	21.15	20.22
Length of hind foot	16.10	10.09	. .	22.50	20.38	21.12	21.05
Spread of antler	9.64	5.64	8.07	. .	6.06	10.54	12.70
Circumference of main beam	10.95	6.69	9.07	11.98	—	13.80	13.22
Length of antler	7.70	2.07	8.70	5.02	2.70	—	10.15
Number of points	12.49	7.09	7.75	12.08	8.07	13.01	—

Part 2: t values for residual regressions for each male deer measurement with effects of indicated measurements removed

Length of hind foot and antler	6.48	3.75	. .	7.70	6.41	—	5.84
Dressed weight and lengths of hind foot and antler	. .	1.27	. .	7.01	4.46	—	7.63

As a straightforward procedure for selection, an intuitive method was tried. This consisted in ranking, in Table III, the values of t in each column of Table II, giving the smallest value the rank 1. A small value of t indicates that the measurement whose effect was removed is closely related with the measurement named at the head of the column. Hence, in Table III, the smallest horizontal total of ranks will indicate the one measurement most closely related, on an average, to all others. The results were consistent with our first interpretation with respect to length of antler and length of hind foot.

To verify our own interpretation of these t values, instead of pro-

ceeding with the development of one function only, we tested several different combinations of measurements, some of which were not promising. Since the 9-month animals were without antlers, two separate

CHART I

AGE DISCRIMINANT FUNCTION FOR MALE DEER USING THE THREE MEASUREMENTS, DRESSED WEIGHT, LENGTH OF HIND FOOT, LENGTH OF ANTLER

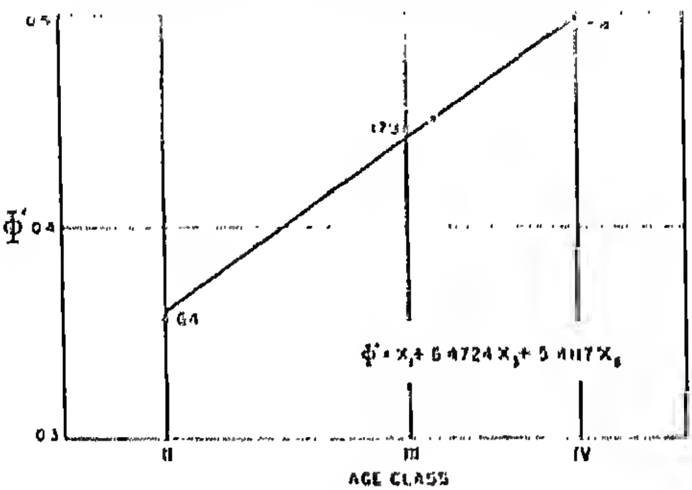


TABLE III

RANK ACCORDING TO EFFECT REMOVED OF THE t VALUES IN TABLE II

Measurement—effect removed	Dressed weight	Length of body	Length of hind foot	Spread of antler	Circumference of main beam	Length of antler	Number of points	Total of ranks for effect removed
Dressed weight	—	1	1	4	4	4	4	18
Length of body	5	—	2	5	5	5	5	27
Length of hind foot	6	6	—	6	6	6	6	36
Spread of antler	2	3	5	—	2	1	2	15
Circumference of main beam	3	4	6	2	—	3	3	21
Length of antler	1	2	4	1	1	—	1	10
Number of points	4	5	3	3	3	2	—	20

analyses were made for the male deer. One, comparable to that for the female deer, included all age groups but only the body measurements; in the other, the 9-month animals were omitted.

To determine the relative effectiveness of these functions in estimating deer age, the procedure used for one measurement was followed in

TABLE IV

PROBABILITIES OF MISCLASSIFYING DEER FOR VARIOUS MEASUREMENTS SINGLY AND IN COMBINATION AS A BASIS FOR ESTIMATING AGE, AND COMPOSITE TEST

Males of age groups, II, III, and IV[1]: Dressed weight	Length of body	Length of hind foot	Spread of antler	Circumference of main beam	Length of antler	Number of points	Probabilities, by age groups: II	III	IV	U	All animals[2]: Dressed weight	Length of body	Length of hind foot	Probabilities, by age groups: I	II	III	IV	U
Male																		
Single measurement																		
•							.22	.48	.26	9.40	•			.05	.27	.47	.20	9.07
	•						.27	.62	.35	6.70		•		.08	.34	.01	.35	8.15
		•					.28	.70	.42	5.27			•	.07	.35	.70	.42	7.08
			•				.18	.45	.27	11.20								
				•			.20	.47	.27	10.96								
					•		.17	.43	.26	11.77								
						•	.20	.51	.31	9.45								
Two measurements																		
•	•						.19	.46	.27	9.80	•	•		.01	.22	.44	.26	10.07
•		•					.18	.46	.28	9.90	•		•	.03	.20	.48	.28	11.34
•			•				.10	.35	.26	13.01		•	•	.04	.27	.58	.35	9.81
•				•			.13	.39	.25	12.28								
•					•		.11	.35	.24	13.00								
	•	•					.23	.54	.33	7.60								
	•			•			.13	.40	.27	11.63								
		•	•				.09	.37	.28	12.87	•	•	•	.03	.20	.46	.28	11.61
		•		•			.12	.39	.27	11.99								
		•			•		.09	.36	.27	12.07								
			•	•			.10	.37	.27	12.59								
			•		•		.10	.37	.27	12.60								
				•	•		.11	.38	.27	12.49								
Female																		
Three measurements																		
•	•	•					.18	.46	.25	10.09	•			.00	.34	.01	.33	8.50
•		•		•			.11	.36	.25	12.65		•		.15	.50	.75	.34	5.48
•		•			•		.09	.34	.25	13.30			•	.13	.58	.81	.30	5.12
•			•		•		.09	.33	.24	13.42								
		•	•		•		.08	.34	.28	13.63								
Four measurements																		
•	•	•		•			.11	.36	.25	13.67	•	•		.06	.35	.01	.32	8.00
•	•		•		•		.09	.33	.24	13.40	•		•	.05	.34	.01	.32	8.08
•		•	•		•		.09	.34	.25	13.60		•	•	.11	.53	.76	.33	6.48
•			•	•	•		.10	.34	.24	13.48								
Five measurements																		
•		•	•	•	•		.08	.33	.25	13.86	•	•	•	.07	.38	.63	.32	8.70

[1] Total number, 397. Class I animals omitted because they have no antlers.

[2] Males, 602; females, 425.

obtaining the mean class values, normal deviates, and probabilities of misclassification. The probabilities for the various discriminant functions considered have been assembled in Table IV. These may be used for an appraisal of the several combinations as indicators of age. The first half of the table gives the results from data for the three upper age classes of males (the deer with antlers), based on 307 animals, while the second half is from data in which all four age classes were represented, 502 male and 425 female deer.

To examine one of the functions graphically, the mean value for each age class was calculated by substituting in it the corresponding measurement means from Table I. Since it is necessary to maintain only the proportionality of the λ coefficients, the calculations may be simplified by assigning to one coefficient the value unity and adjusting the others by division. One of the functions and the means are shown in Chart I plotted against the age-group values, which are maintained on a logarithmic scale. The linear representation may then be passed through the weighted means with a slope equal to the derivative of the function. Since B_p was the regression of x_p on age, the derivative is for all practical purposes equal to $\lambda_1 B_1 + \lambda_2 B_2 + \cdots$.

COMPOSITE TEST FOR RANKING DISCRIMINANT FUNCTIONS

To judge two or more discriminant functions for comparative effectiveness, it was desirable to use a single value rather than a series of probabilities. The best discrimination between two groups would be accomplished with the function that produced the largest difference between the calculated group means. When more than two groups are under consideration, an average of all the differences between successive groups is given by the slope of the function which represents the change in the function per unit change in the independent variable. Thus, for the purpose of discrimination, the maximum slope is to be desired. This, of course, was the original intent in maximizing the regression sum of squares in calculating the discriminant function. We might say, then, that of two discriminant functions the one with the greater slope is the more effective. However, the variability within classes must be taken into account if a classification of individuals is to follow. Then the change in standard units per unit change in age might be considered a measure of the discriminating effectiveness of any particular function. This is the ratio of the slope to the standard error of the function; and the larger this ratio is, the better does the compound measurement discriminate among age groups.

If two functions have equal slopes, the one with the smaller error

should be considered the more effective. This is consistent with the notion just expressed that the larger the ratio, the better is the discrimination. Therefore, we propose the intuitive test, by definition,

$$U = \frac{\text{slope of the function}}{s_\phi}$$

where the slope of the function is the derivative of ϕ and is $\sum \lambda_p B_p$. The U values are listed in Table IV for all combinations shown.

In the three-age analysis no single measurement discriminates age with any acceptable degree of precision, the failure being particularly striking for animals of age class III. Of the body measurements, dressed weight is considerably better than either of the other two. Adding one or both of the length measurements does not improve the precision of estimate perceptibly. A function of these two without dressed weight is the poorest for any combination of two.

Any one of the antler measurements, except number of points, is better than dressed weight. Length of antler is the best single measure; number of points has small value. Compounds of two antler measurements give better estimates than one alone, spread and length of antler being the best. A slightly better compound of two is one body with one antler character, preferably length or spread of antler. From the U value, the compound length of hind foot and length of antler would seem to be as effective as any including dressed weight.

The best combination of three is length of hind foot, spread and length of antler. When one body and two antler measurements are combined, no advantage is gained by including another antler measurement. Compare, for example, the values obtained by combining dressed weight and spread and length of antler with those obtained by combining these three with circumference of main beam.

We may now check the interpretation of the ranking of the t tests shown in Tables II and III. Length of antler has the smallest total of ranks, and has the largest U value. Length of antler and dressed weight were first and third in the ranking and their combination has the large U value of 13.00. This barely differs from the value 13.01 for spread of antler and dressed weight, which have rankings 2 and 3, respectively.

The highest U value for three measurements is obtained with length of antler, spread of antler, and length of hind foot, which ranked first, second, and last, respectively. It was a satisfaction to find that our intuitive selection of the hind foot was upheld when at least three measurements were involved.

The three measurements just mentioned together with dressed weight formed the best combination of four, as is judged by $U = 13.80$. These measurements, as we have seen, had the ranks 1, 2, 3, and 7.

In dealing with overmature deer the chance of misclassifying is only slightly less when a compound is used instead of any one single measurement other than length of body, length of hind foot, or number of points. The greatest advantage of the discriminant function over a single character is in age groups II and III of male deer. The quantity U or its square, may prove of value, since it is independent of any proportional shift in the λ's.

After this paper was prepared, we ran across a discussion of a similar, if not identical measure, for the comparison of two discriminant functions. This was done by David Durand in his book *Risk Elements in Consumer Instalment Financing* (3). He confined his treatment to the comparison of two groups only, with application to bad and good loans. Strangely enough he too had called his function U. It is equal to the $(\Sigma l_i a_i / \sqrt{\Sigma l_i l_j s_{ij}})$ where a_i is the difference between the means of the two groups for the ith measurement and the l's are the coefficients.

It is seen that our U is a generalization of Durand's in the case of several groups, for terms $\lambda_i B_i$ in our numerator, which was obtained from the slope of the function, are such generalizations of the terms $l_i a_i$ in Durand's. The denominator, which we felt was needed to tame down the U value in the case of great variability, is identical with his.

It might not prove a difficult matter to generalize the analysis-of-variance test as outlined by Fisher (5, p. 377) for testing the significance of a single discriminant function. As previously stated it has been done for large samples. However, we are not so much interested in the significance of a single function as we are in selecting the best from among a number of such similar functions. It would seem reasonable to suppose that the distribution of the generalized U would be found to be related to that of some generalized analysis-of-variance test as an extension of the distribution of Hotelling's T^2. It might be added here that the comparison of two discriminant functions undoubtedly involves the same difficulties as those connected with comparisons of multiple regressions. These have been discussed by both Fisher (7) and Hotelling (8).

APPLICATION OF DISCRIMINANT FUNCTIONS TO DATA NOT USED IN DEVELOPING THE FUNCTIONS

It remains to show how the results from this study may be applied to field data to determine age, although it cannot be too strongly empha-

sized that the functions derived and the conclusions reached for one herd of Whitetail deer are applicable only to data from similar herds. In addition to the data used in computing the functions, we had 407 sets of measurements with which we could test those functions not involving length of body and length of hind foot. The compound shown in Table V was applied to the measurements of the individual deer and age group limits were determined as for the single measurement. The theoretical probability of misclassification may now be compared with the proportion of deer actually misclassified in each age group. With these four measurements, calculated ratios for the animals of ages II, III, and IV are .16, .39, and .26, respectively—all slightly larger than the theoretical values, .10, .34, and .24. With the three measurements, dressed weight, spread and length of antler, the probabilities are .16, .38, and .33—again all larger than the theoretical values, .09, .33, and .24 (Table IV). With the two antler measurements only, similar differences were observed.

The totals in the last column of Table V show the number of animals that would be classed in the three age groups by use of the function alone. These frequencies may be compared with the true numbers for the three groups, 76, 261, and 70. With all functions, the number of mature deer is underestimated in favor of the two other groups.

TABLE V

APPLICATION OF THE CALCULATED DISCRIMINANT FUNCTION, $\Phi = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_p x_p$, FOR DETERMINING THE AGE OF A GROUP OF MALE DEER NOT USED IN DEVELOPING THE FUNCTION

$\Phi' = x_1 + 2.3321 x_2 + 5.2760 x_3 + 2.0728 x_4$

Theoretical probabilities of misclassification: .10, .34, .24

Theoretical age group based on Φ'	Age class II		Age class III		Age class IV		Total (theoretical age)
	Number	*Per cent*	*Number*	*Per cent*	*Number*	*Per cent*	
IV 332.7 and above	0	0.0	49	18.8	52	74.2	101
III 249.9–332.6	12	15.8	160	61.3	16	22.9	188
II 249.8 and below	64	84.2	52	19.9	2	2.9	118
Total (dental age)	76	100.0	261	100.0	70	100.0	407

FURTHER INVESTIGATIONS NEEDED

In order to standardize the discriminant function method, it would be necessary to carry out three additional investigations. One should be made on a method of selecting the best of several measurements. If the t test is used, the ranking scheme should be investigated further. Another should be made on the validity of the test of significance of the

lambda coefficients, and a third for determining a test of significance of the difference between two discriminant functions in the case of more than two groups. Perhaps the distribution of the composite test, U, will be adequate for this if it is found to be some generalization of that of Hotelling's T^2.

LITERATURE CITED

1. M. M. Barnard. "The Secular Variations of Skull Characters in Four Series of Egyptian Skulls," *Ann. Eugenics* 6 (1935): 352–371.
2. Gertrude M. Cox and William P. Martin. "Use of a Discriminant Function for Differentiating Soils with Different Azotobacter Populations." *Iowa State Col. Jour. Sci.* 9 (1937): 323–331.
3. David Durand. *Risk Elements in Consumer Instalment Financing.* Natl. Bur. Econ. Res., Inc., Studies in Consumer Instalment Financing, No. 8 (1941).
4. R. A. Fisher. "The Use of Multiple Measurements in Taxonomic Problems." *Ann. Eugenics* 7 (1936): 179–188.
5. ———. "The Statistical Utilization of Multiple Measurements." *Ann. Eugenics* 8 (1938): 376–386.
6. ———. *Statistical Methods for Research Workers.* Ed. 7, rev. and enl. 1938.
7. ———. "The Precision of Discriminant Functions." *Ann. Eugenics* 10 (1940): 422–429.
8. Harold Hotelling. "The Selection of Variates for Use in Prediction with Some Comments on the Problem of Nuisance Parameters." *Ann. Math. Statis.* 11 (1940): 271–283.
9. P. L. Hsu. "On the Problem of Rank and the Limiting Distribution of Fisher's Test Function." *Ann. Eugenics* 11 (1941): 39–41.
10. Khint Maung. "Discriminant Analysis of Tocher's Eye-Colour Data for Scottish School Children." *Ann. Eugenics* 11 (1941): 64–76.
11. Barry C. Park and Bessie B. Day. *A Simplified Method for Determining the Condition of Whitetail Deer Herds.* U. S. Dept. Agric. Tech. Bul. 840 (1942).
12. H. Fairfield Smith. "A Discriminant Function for Plant Selection." *Ann. Eugenics* 7 (1936): 240–250.
13. Noel Wallace and R. M. W. Travers. "A Psychometric Sociological Study of a Group of Specialty Salesmen." *Ann. Eugenics* 8 (1938): 266–302.

THE NEW WHEAT CONDITION FIGURES, BASED ON WEATHER FACTORS, FOR THE PRAIRIE PROVINCES*

By Fred H. Sanderson
Harvard University

In June, 1941, the Bureau of Statistics of the Dominion of Canada began to issue for the Prairie Provinces wheat condition figures based on weather factors. At the same time it discontinued the collection of the reports of crop correspondents for this crop. As far as is known, this is the first time that a government agency has attempted to base its official crop reports exclusively upon objective measurements of the meteorological conditions prevailing prior to and during the growing season.

This change is the third that has been effected during the comparatively short history of the condition reports issued by the Canadian crop reporting service. Up to 1918, the condition as reported by crop correspondents was expressed as a percentage of a "full" or "standard" crop, a definition similar to the one adopted by the United States Department of Agriculture. In 1919, the Dominion Bureau of Statistics adopted a presumably more precise method whereby the correspondents were asked to report condition as a percentage of the average yield of the previous ten years. From 1931 to 1940, the crop condition of 100 per cent was interpreted to be equal to the "long-time average yield per acre." The latter was first defined as the average yield for 1908 to 1930; with the recent shift to condition figures based on weather factors, the base was changed to the 1908–1940 average yield.

Where crop condition figures are based on the returns of crop correspondents, frequent changes in definition are unfortunate for several reasons. Condition figures as reported by correspondents are *opinions*, based on the appearance of the growing crop and supplemented by the farmer's experience of the general effect of the weather on yield. Like all opinions, they are affected by certain elements of psychological *bias*. The 100 per cent condition may be *defined* as a "condition promising a yield equal to the long-time average." But it does not usually follow that in practice, a 100 per cent condition is associated, on the average, with the average yield; a 50 per cent condition with one-half of the average yield, and so forth; or that a crop condition defined by 100 is equal to four times a crop condition defined by 25.

* The study upon which this paper is based was made while the author was a research associate at Harvard University and a part-time employee of the Agricultural Marketing Service, United States Department of Agriculture; it was the outgrowth of a weather-yield research project for improving statistical methodology in forecasting crop yields. The author is indebted to Mr. Joseph A. Becker, of the United States Bureau of Agricultural Economics, for reading and criticizing the manuscript.

Generally speaking, crop correspondents in reporting condition show a triple bias: (1) In underestimating average yield prospects, (2) In underestimating the intensity of the year-to-year variations in yield, (3) In shying away from reporting more than a "fair" yield prospect at any time.

The tendency of crop reporters to underestimate average yield prospects is relevant only where the condition figures are defined as percentages of some definite "average" or "normal" yield. It will then appear that a reported condition of 100 would, on the average, actually indicate a probable yield considerably exceeding the "average" or "normal." If the final yield is expressed as a percentage of the long-time average and plotted along the y-axis of a system of coordinates, and condition along the x-axis, the line of average relationship would not be the diagonal defined by the points 0:0, 50:50, 100:100, etc., but would lie above this line. This downward bias of the condition reports would increase as the season advances.[1]

If the condition figures provided a correct indication of the *intensity* of the variations in yield, an increase in the condition by a given percentage would indicate, on the average, an equal *relative* increase in the yield. Hence, if condition is plotted along the horizontal and yield along the vertical axis, both being expressed as percentages of their respective means, and by using the same scale along both axes, the regression coefficient would be equal to unity; or, in other words, the slope of the regression line or line of average relationship would be 45 degrees.

For most condition-yield relationships, however, an increase of ten per cent in "condition" would indicate an increase in yield of *more* than 10 per cent, and the line of regression would meet the x-axis at an angle of more than 45 degrees.

But in addition, this bias is not constant over the whole range of prospective yields. Crop reporters would hesitate to report more than a

[1] In Saskatchewan, it was found that a condition of "100 per cent" reported at June 30 would correspond, on the average, to a 124 per cent yield (24 per cent above the average 1921–1937), whereas the same "condition" reported on July 31 would indicate a probable yield of 137 per cent. In Alberta, the corresponding figures are 109 and 121 per cent, respectively. Consequently, in comparing condition as reported, for example, at May 31 and June 30, allowance should be made for this tendency of condition figures to decline as the season advances. Such statements as the following, quoted from the official crop report of the Canadian government of July 6, 1940, may serve to illustrate this point:

"... slight declines in the condition of the wheat crops in Manitoba and Alberta were experienced during June while a more serious loss in condition was suffered in Saskatchewan. Expressed in percentages of the provincial long-time average yields per acre, Manitoba and Alberta wheat conditions both declined from 98 to 96 during June. ... In Saskatchewan, the wheat condition declined 5 points from 94 at May 31, this year, to 89 at June 30."

The report does not indicate that there is normally a decline in the condition figures from May 31 to June 30. The average "condition" as of May 31 during the period 1921–1937 was 94 for Saskatchewan and 98 for Alberta. The corresponding figures for June 30 are 87 and 93. Thus, in 1940, the condition actually *improved* from average to two above average in Saskatchewan, and from average to three above average in Alberta. The apparent decline in the condition figure has no real significance.

"slightly better than average" yield prospect at any time. In other words, reported condition would not increase proportionately with prospective yield. Condition tends to be a very sensitive indicator of low yield prospects, but becomes less sensitive as prospects improve.[2]

Elements of bias in condition figures similar to those described above have been observed in many countries where crop forecasts are based on numerical indications regarding the appearance of the growing crop as gathered from crop correspondents, and adequate methods to deal with this problem have been worked out.[3] The success of these methods depends to a great extent on the stability of the psychological factors involved, and, in particular, on the homogeneity of the conditions surrounding the collection of the reports. Any change in the procedure of collecting the reports or in definition may constitute a break in homogeneity and introduce unpredictable changes in the relationship between reported condition and yield. A long series of observations under homogeneous conditions is needed to eliminate bias from the reported condition figures. Any radical innovation will therefore impair the usefulness of a series of condition reports for forecasting purposes.

In a large number of countries, including the United States, condition figures as reported by crop correspondents were found to be valuable indicators of prospective yields; they have been used sometimes with remarkable success as a basis of forecasting the yield several months prior to the harvest. The Dominion Bureau of Statistics, however, came

[2] Take, for example, the case of Saskatchewan, June 30. When yield prospects were poor, an increase from 8 bushels below average to 6 bushels below average was accompanied, on the average, by an increase in "condition" as reported by crop correspondents from −53 to −16, or 37 points. An identical increase in yield when yield prospects were favorable (plus 5 to plus 7 bushels) was accompanied by an increase in reported condition from plus 14.5 to plus 16.8, or 2.3 points.

Where condition is plotted against yield per acre *harvested* (a practice resorted to when data for acreage sown are unavailable), it may be argued that the departure from linearity might possibly be due to abandonment of acreage subsequent to the condition report. It is, indeed, possible that the condition reports made by crop correspondents—particularly the earlier reports—refer more closely to acreage sown than to acreage harvested. Widespread abandonment of acreage subsequent to the report will, therefore, introduce an apparent downward bias in the condition figures when the latter are compared with final yield per *harvested* acre. As abandonment of acreage tends to be associated with low yields rather than with high yields, this apparent downward bias will be particularly noticeable in years of poor yields, while in years of above-average yields and correspondingly small extent of abandonment, it would scarcely affect condition figures at all.

If this view was more than a partial explanation of the curvilinear bias present in condition figures, the relationship between reported condition and yield per acre *sown* would be a straight line. It was found, however, that where yield data were available per seeded acre as well as per harvested acre, the curvature was somewhat attenuated in the former case, but did not disappear. A further confirmation of this is provided by the case at hand, for the Canadian yield data refer to *acreage sown*.

However, straight-line relationships between "condition" and yield may be found, especially early and late in the season. In the early part of the growing season, the total correlation between "condition" and yield is usually poor, and the curvilinearity of the relationship is frequently blurred by errors. As the season advances, the departure from linearity would become more apparent. Shortly before harvest, however, prospects of bumper crops begin to become recognized, and the relationship between "condition" and yield tends to approach a straight line.

[3] See United States Department of Agriculture, Miscellaneous Publication No. 171, *The Crop and Livestock Reporting Service of the United States*, Washington, D. C., November 1933, pp. 22–32.

to the conclusion that the "condition figures obtained during the growing season are not sufficiently precise to serve as a basis for quantitative forecasting."[4]

No forecasts of production have been issued by the Canadian government since 1927. Up to 1918, official forecasts of yield and production obtained by the "par" system adopted by the United States Department of Agriculture in 1912 had been issued. The condition as reported by crop correspondents was expressed as a percentage of a "full" or "standard" crop. Forecasts of production were issued early in June, July, and August.

In 1919, at the suggestion of the International Institute of Agriculture, the Dominion Bureau of Statistics shifted to a presumably more exact method whereby the correspondents were asked to report condition as a percentage of a ten-year average. In making a forecast of yield, this reported percentage was applied directly to the average yield of the previous ten years. The reasons given for this change are as follows:

"The disadvantage of (the old) system is that the 'standard' or 'full' crop is a matter of opinion and dependent entirely upon the mental conception of each correspondent as to what such a standard should be. Consequently, the system has no true statistical basis."[5]

Within a year, this method proved unsatisfactory, and it was discontinued, though without mention that the true relation between condition figures as reported and final yields might differ from the one presupposed by the definition officially adopted.[6]

Between 1920 and 1923, the Dominion Bureau of Statistics did not publish forecasts of yield and production as of June 30. Forecasts on the basis of June 30 condition were resumed in 1924, but definitely terminated in 1927. Forecasts using condition at July 31 were continued until 1927. Since 1928, no official interpretation of condition figures has been made by the Canadian government. A preliminary estimate of production, issued early in September, is based on crop correspondents' estimates of probable yield.

Do the condition figures for spring wheat in the Prairie Provinces have value for forecasting yield per acre? To investigate this question, the condition figures as reported by correspondents during the period 1921–1937 were plotted against the final yields, separately for the two

[4] C. F. Wilson, Dominion Bureau of Statistics, Ottawa: *Relations Between Weather Factors and Wheat Yields in Western Canada.* Paper presented at the joint meeting of the Canadian Society of Technical Agriculturalists and the American Statistical Association, Ottawa, June 29, 1938. Mimeographed.

[5] Dominion of Canada, Department of Trade and Commerce, Census and Statistical Office: *Monthly Bulletin of Agricultural Statistics,* Vol. 11, No. 116, April 1918, p. 103.

[6] Canada, Dominion Bureau of Statistics: *Monthly Bulletin of Agricultural Statistics,* Vol. 13, No. 143, July 1920, p. 169.

provinces of Saskatchewan and Alberta. The significant correlation coefficients or indexes[7] of Table I were found.

TABLE I

Date of report	*Saskatchewan*	*Alberta*
May 31	0.284	not significant
June 30	0.570*	0.525*
July 31	0.839*	0.614*

* Indicates curvilinear relationship.

In view of the comparatively high correlations[8] obtained, it would seem that the Canadian condition figures as reported by crop correspondents can be used successfully as a basis of forecasting wheat production.

When the Dominion Bureau of Statistics made its recent shift from condition figures based on crop reporters' returns to condition figures based on weather factors, it announced the change with the following statement:

[7] The square of the correlation coefficient (or index) is used rather than the index itself. The squared correlation coefficient or index measures the percentage of the total year-to-year variability in yields that is associated with the predictor (or predictors).

All correlation coefficients and indexes have been corrected for bias, according to the formula

$$\bar{R}^2 = 1 - \left[(1 - R^2) \cdot \frac{n-1}{n-m} \right]$$

where n is the total number of observations and m the number of constants used. See M. Ezekiel: *Methods of Correlation Analysis*, 1930, p. 177.

These correlations were computed in 1938. Forecasts based on condition reports for 1938 and 1939 were well within the range of the errors of estimate during the period covered by the analysis. Of ten forecasts, six did not exceed the standard error of estimate; none exceeded twice the standard error. In 1940, however, the condition figures, with two exceptions, fell rather far off the mark. Throughout the Prairie Provinces, the 1940 condition figures were unusually poor indicators of the final yield. In some cases, the deviations from the line of average relationship were among the largest hitherto observed.

However, none of these deviations was larger than 1.7σ. In other words, even the largest of these errors did not exceed what one would expect to find in random sampling, at least once in about 10 cases. Moreover, in making this statement, we assume that our statistical regression curve is identical with the "true" regression. If the standard error of the regression is taken into account, we obtain, for the individual forecast, a standard error which is appreciably larger than the standard error of estimate (compare M. Ezekiel: *Methods of Correlation Analysis*, 2nd edition, 1941, p. 344).

[8] As it has been pointed out above, the definition of the 100 per cent condition was changed in 1931 from "the condition promising the average yield per acre of the previous 10 years" to "a condition promising the average yield 1908–1930."

Although strictly speaking there is no safe basis for any guess regarding the effect of this change in definition, there is a considerable question whether the slight modification may not have been overlooked entirely by most crop correspondents. In computing the above correlations, therefore, there seemed to be justification for neglecting it altogether. In fact, experience in various countries has shown that whatever the basic definition adopted by the crop reporting service, the base or 100 per cent condition in the minds of crop correspondents turns out to be a fairly stable concept, corresponding approximately, not to the long-time average, but to a long-time "normal" condition which would indicate above-average yield prospects. Crop correspondents seem to have a definite idea as to what their fields "ought" to produce in the absence of damage. In this sense, a "normal" crop is taken to be a full crop, and years of widespread or complete crop failure appear to be largely forgotten when the current crop is appraised in terms of the "normal" condition. The Canadian condition figures are no exception to this rule.

In Table 3 a comparison is afforded between the wheat condition figures in the Prairie Provinces based on the weather factors, and those based on the returns of crop correspondents as they both relate to the condition figures corresponding with the *final yields per acre for each year's crop.* The three sets of condition figures employed in the comparison have been adjusted to percentages of the new 1908–1940 long-time average yields per acre for each province.

A comparison of the condition figures shown indicates that on an average of 7 out of every 10 instances, the condition figures based on weather factors approximate more closely the final yield of the crop than have the previously published condition figures. In the majority of the remaining instances, where the previously employed condition figures have been more sensitive to changes in crop prospects, the reasons have been the occurrence of heavy rust or insect damage, which in turn have not been related to changes in the weather factors. In actual practice in the future, the wheat condition figures based upon the weather factors can be adjusted in the event of abnormal grasshopper activity, or in the event of rust damage, which will be less probable in the future than in the past. Accordingly, use of the condition figures based on weather factors is expected year in and year out to provide a more accurate indication of the numerical change in Prairie wheat prospects than has been available in the past.[4]

This statement fails to do justice to the usefulness of the condition figures based on the returns of crop correspondents and it is apt to give rise to false hopes as far as the new system is concerned. In fact, the above comparison seems misleading on a number of counts which cumulatively contribute to distort the picture in favor of the new system.

(1) In attempting to improve the condition figures as reported by crop correspondents, the latter were apparently adjusted to the 1908–1940 base in the following manner:

1921–1930:

$$\text{Adjusted cond.} = \text{published cond.} \times \frac{\text{Av. yield of preceding 10 years}}{\text{Av. yield 1908–1940}}$$

1931–1940:

$$\text{Adjusted cond.} = \text{published cond.} \times \frac{\text{Av. yield 1908–1930}}{\text{Av. yield 1908–1940}}.$$

It was assumed that a change in the base would, on the average, lead to a proportionate change in the condition figures. Owing to the rather complex relationship existing between "condition" and yield, it turned out that this "adjustment" of the condition figures, instead of improving the correlation with final yield, actually resulted in lower coefficients in all cases. In Table II the correlation coefficients or indexes obtained with the unadjusted condition figures are compared with the ones obtained after adjustment.

[4] Canada, Dominion Bureau of Statistics, Agricultural Branch: *Quarterly Bulletin of Agricultural Statistics*, Vol. 34, No. 390, April–June 1941, pp. 115–119.

TABLE II

Date of report	Saskatchewan		Alberta	
	Before adj.	After adj.	Before adj.	After adj.
May 31	0.281	0.000	0.000	0.000
June 30	0.670*	0.320*	0.525*	0.505
July 31	0.839*	0.771*	0.614*	0.557*

* Indicates curvilinear relationship. Regression curves were fitted wherever this led to a statistically "significant" increase in the correlation.

The corresponding scatter diagrams appear to corroborate the hypothesis that the slight change in definition in 1931 might well have passed practically unnoticed by the crop correspondents.

(2) In the comparison made by the Dominion Bureau of Statistics of the two sets of condition figures, account was not taken of the fact, indicated above, that wherever condition figures are defined as percentages of some definite "average" or "normal" condition, they tend to fall short of actual wheat prospects.

(3) Allowance was not made for the tendency of crop correspondents—also described above—to underestimate the intensity of variations in yield.

(4) Allowance was not made for curvilinear bias present in most condition figures based on crop correspondents' returns.

Ordinary correlation methods such as are used in the United States make allowance for these systematic biases.

In the series of "condition figures" calculated from weather factors with which comparison is made, the types of bias summarized under headings (2) to (4) are not present because: (a) The method of multiple correlation starts out by equalizing the average computed yield with the average actual yield.[19] (b) The least-squares principle automatically allows for differences in the intensity of the variations in the predictor series as compared with those present in the dependent variable. The formula is so constructed that if computed yields are compared with actual yields, the positive and the negative errors just balance each other; the sum of the errors of estimation is equal to zero. (c) Wherever the relationship between weather factors and yield could not be represented adequately by a straight line, curvilinear correlation methods were applied by the Dominion Bureau of Statistics in working out the weather-crop formulae.

In order to obtain a fair comparison between the two methods, both sets of condition figures must be analyzed in the same way. In the comparison made by the Bureau, only the weather factors got the benefit

[19] Compare, however, footnote 12, second paragraph.

of an adequate statistical treatment. When correlation methods are applied in both cases, it appears that for Saskatchewan the three correlation coefficients (or indexes) computed on the basis of the old-style condition figures are all higher than those obtained with the new figures based on weather factors. In Alberta the correlations obtained with the new system are not materially better than those found on the basis of the old method.[11] The coefficients[12] (or indexes) obtained in the two cases are shown in Table III.

TABLE III

Date of report	Saskatchewan			Alberta		
	Cond. as published (Based on crop reporters' returns)	Cond. based on weather factors		Cond. as published (Based on crop reporters' returns)	Cond. based on weather factors	
	1921–1937	1921–1937	1921–1940	1921–1937	1921–1937	1921–1940
May 31	0.284	0.261	0.141	0.090	0.452	0.363
June 30	0.670	0.499	0.559	0.525	0.721	0.688
July 31	0.839	0.800	0.789	0.611	0.745	0.755

(5) A recent study[13] by the writer indicates that it is possible to improve the wheat condition figures as reported by crop correspondents for the provinces of Saskatchewan and Alberta, with the aid of one or two weather factors which crop correspondents apparently do not sufficiently consider when judging the condition of the crop. It was found

[11] Saskatchewan, the most important of the wheat-producing provinces, accounts for more than one-half of the Canadian wheat acreage; Alberta for one-third.

[12] The correlation indexes computed for "condition based on weather factors" differ from those shown in Dr. Wilson's study ("The Influence of Precipitation and Temperature on Wheat Yields in the Prairie Provinces, 1921–1940," Canada, Dominion Bureau of Statistics, *Quarterly Bulletin of Agricultural Statistics*, Vol. 34, No. 391, July-September, 1941, pp. 167–187. Cf. especially p. 187). The latter were computed after exclusion of years in which rust seriously reduced yields and, for Saskatchewan, after adjustment for damage by entomological factors in the years 1931–1939. In an appraisal of the adequacy of the weather-crop formulae developed by the Dominion Bureau of Statistics, however, errors due to "extraneous influences" must be taken into full account, since the problem of predicting the effect of these factors on yield is a particularly difficult one.

If, however, the regressions derived from the reduced number of observations are applied to the *whole* period, a discrepancy appears between the average estimated yield and the average actual yield for 1921–1940 (cf. *loc. cit.*, Table 2, cols. 7 and 8, and Table 8, cols. 7 and 8). Adjustment was made here for this discrepancy by computing linear correlations between estimated and actual yields. The correlation coefficients thus obtained were then corrected to allow for degrees of *freedom*. The latter were derived by subtracting the number of constants indicated in the study (compare also Table IV) from the total number of observations.

The following factors were used in Dr. Wilson's study for Saskatchewan, July 31: Seven years' weighted average precipitation, preseasonal precipitation, spring precipitation, spring temperature, summer precipitation, summer temperature. Curvilinear functions have been fitted to four of these factors.

[13] Fred H. Sanderson, Bureau of Agricultural Economics, United States Department of Agriculture: *The Use of Condition Reports and Weather Data in Forecasting the Yield per Acre of Wheat.* (Mimeographed) October 1942.

that forecasts based on official condition reports may be significantly improved by allowing for variations in subsoil moisture and for weather conditions prevailing during the period immediately preceding the report.

Lack of subsoil moisture would not necessarily affect the appearance of the crop early in the season, provided spring rains had been sufficient to germinate the crop and give it a good start. In fact, lack of subsoil moisture might not greatly affect the appearance of the crop until after heading time or later when the plant requirements for moisture are at or near their maximum, and the moisture due to precipitation during the growing season is nearly or completely used up by the plant. One might infer, therefore, that the condition of the crop in the early part of the season is influenced largely by current precipitation, and the reserve of moisture in the lower levels of the soil is neglected by crop correspondents when estimating condition. Unfortunately, no direct data regarding subsoil moisture were available for this study, and recourse had to be taken to the indirect representation of variations in soil moisture by one of its determinant factors, namely the precipitation received during the months of September, October, and November of the year preceding harvest.

It appeared likely, furthermore, that crop correspondents would fail to allow for the effect of weather conditions immediately preceding the report. As Barnes and Hopkins[14] have pointed out, even a heavy downpour does not lead to a considerable increase in soil moisture at the depth available for plant use until ten days or more after its occurrence;[15] that is, the appearance of the crop may not be affected immediately. One would expect that this lagged effect of rainfall is particularly important during the period preceding heading time (about the end of June and the beginning of July) which is considered a critical period with regard to wheat yields in the Prairie region. Consequently, it is particularly at this time of the year that any tendency of crop correspondents to neglect this factor would result in a considerable loss of reliability and timeliness in their reports.

These two hypotheses were tested statistically and it was found that the increase in the correlation coefficients or indexes resulting from the inclusion of these factors is highly significant. It would seem, in fact, that in Alberta condition taken in conjunction with the above weather factors would provide a better basis for forecasting wheat yields early in July than was possible a full month later on the basis of condition

[14] S. Barnes and E. S. Hopkins: *Soil Moisture and Crop Production*, Dominion of Canada, Department of Agriculture, Bulletin 130 — New Series, Ottawa, 1930.

[15] Work with soil tensiometers at Ames (Iowa) (1938) showed that it took a week to ten days for a precipitation of two inches to penetrate 24 inches.

alone. In Saskatchewan, forecasts made early in July by means of the condition-and-weather formulae may be expected to be about as reliable as forecasts made early in August based on condition alone.

In Table IV, the corrected squared correlation coefficients (or indexes) involving condition and preseasonal precipitation are shown for

TABLE IV

Date of report	Variables used	Saskatchewan: Number of constants	Saskatchewan: Squared correlation coef. or index, Uncorrected	Saskatchewan: Squared correlation coef. or index, Corrected*	Alberta: Number of constants	Alberta: Squared correlation coef. or index, Uncorrected	Alberta: Squared correlation coef. or index, Corrected*
May 31	Condition as published	2	0.329	0.281	2	0.019	0.000
	"Adjusted" condition	2	0.054	0.000	3	0.031	0.000
	Condition and preseasonal precipitation	3	0.621	0.567	3	0.318	0.221
	Various weather factors (Wilson)						
	1921–1937	7	0.634	0.284	6	0.623	0.452
	1921–1940	7	0.412	0.141	6	0.623	0.363
June 30	Condition as published	3	0.624	0.570	3	0.584	0.526
	"Adjusted" condition	3	0.405	0.320	2	0.536	0.505
	Condition and preseasonal precipitation	4	0.798	0.751	4	0.702	0.633
	Condition, preseasonal precipitation, 10 days' precipitation and 10 days' temperature	6	0.907	0.865	7	0.916	0.860
	Various weather factors (Wilson)						
	1921–1937	11	0.812	0.499	8	0.844	0.723
	1921–1940	11	0.791	0.550	8	0.863	0.698
July 31	Condition as published	3	0.851	0.830	3	0.662	0.614
	"Adjusted" condition	3	0.800	0.771	3	0.612	0.657
	Condition and preseasonal precipitation	4	0.930	0.914	3	0.727	0.688
	Various weather factors (Wilson)						
	1921–1937	11	0.925	0.800	11	0.897	0.725
	1921–1940	11	0.900	0.789	11	0.884	0.755

* Corrected for number of degrees of freedom. Compare Footnote 7.

the two provinces of Saskatchewan and Alberta, for May 31, June 30, and July 31. Another index, involving, in addition, weather conditions during the ten days immediately preceding the report, is shown for June 30. It would seem that with the exception of the May 31 and July 31 reports for Alberta, the reported condition figures, when treated by correlation methods and supplemented by weather factors, may be expected to afford a better basis of forecasting wheat yields than the weather studies of the Dominion Bureau of Statistics.

(6) In comparing the new condition figures "based on weather factors" with the reported condition figures, it should be noted that the former are based on a least-squares analysis involving a considerable number of variables, whereas the latter, if treated by correlation methods, involve only two or three constants. It may be argued that the conventional correction for degrees of freedom takes care of this. There can be no doubt, however, that as the number of constants used in the analysis increases, the regression formulae derived by least-squares procedures tend to become subject to an increasing extent to the limitations and pitfalls inherent in our statistical technique. The significance of forecasting formulae tends to be reduced by such factors as the presence of joint relationships, inadequately dealt with by the conventional additive methods; the presence of serial correlation; and the appearance of spurious relationships which occur when the same body of data are used to establish a hypothesis and to prove it.[16]

This writer is well aware of the shortcomings of traditional crop reporting and shares the hope that condition figures based on correspondents' reports will be supplemented to an increasing extent or even replaced by objective methods of crop forecasting. It is in this way that progress will be made in developing the techniques for forecasting production. Dr. Wilson's analyses of the influence of weather factors on wheat yields in the Prairie Provinces represent a valuable contribution to this end. Unfortunately, the discontinuance of the collection of wheat condition reports from crop correspondents in the Prairie Provinces renders impossible any direct comparison of the relative merits of the two approaches in future years.

A REPLY

BY C. F. WILSON
Dominion Bureau of Statistics

MR. SANDERSON'S dominant theme is the challenging of a decision made by the Dominion Bureau of Statistics in 1941 to change the basis of its Prairie Province wheat condition figures from the one derived directly from crop correspondents' returns to a new one resulting from a study of the relations between weather and yields.[1] He asserts that the stated reasons for this decision seem "misleading on a number of counts which cumulatively contribute to distort the picture in favor of the new system." He then proceeds to elaborate six counts to which in turn this reply is addressed.

[16] In Dr. Wilson's study—at least in the case of the Manitoba and Alberta analyses—the choice and combination of the variables were guided to a considerable extent by "inspection of the scatters" (*op. cit.*, pp. 177, 181).

[1] Mr. Sanderson's footnote 12.

(1) The announcement of the change in bases carried in the June 6, 1941, crop report[2] was written for the particular group of people in the various government agencies and grain-handling organizations, in whose plans and decisions condition figures are important. By and large, by virtue of their different training, the users of condition figures are not responsive to statistical refinements. When the Dominion Bureau of Statistics decided that the weather-yield analysis provided a sufficient basis for replacing the unadjusted correspondents' condition figures with which dissatisfaction had become general, there still remained the problem of presenting the new series to its users.

The simplest procedure was to compare the correspondents' condition figures and the weather-yield results, logically those for July 31, with the final yields per acre. To facilitate direct comparisons, all three series had to be adjusted to a common base. Since the new condition figures from the weather-yield analysis were to be published in terms of the 1908–1940 base, the latter was selected. Mr. Sanderson has shown that his regressions of yield on correspondents' condition suffer some impairment when the latter are shifted to the 1908–1940 base. The writer was not, as Mr. Sanderson feels he should have been, comparing the weather-yield condition figures with the results of such regressions. The merits of the latter will be discussed under the three subsequent counts. Meanwhile, to avoid any suggestion of having affected the comparison to the disadvantage of the correspondents' condition figures by shifting them to the 1908–1940 base, direct comparison of the three series can just as readily be made in terms of 1908–1930 base.[3] In this way the condition figures *as published*[4] may be compared directly with the condition based on weather factors and with final yield, when the latter two series are adjusted to the 1908–1930 base. When such comparison is made the results if anything show an improvement over the stated 7:3 ratio in favor of the weather-yield analysis at July 31. It, therefore, cannot be implied that the users of condition figures were misled by the comparison presented between the results of the weather yield study and those of the condition figures hitherto employed.

(2), (3) and (4). These counts are like in kind in that they charge failure to take into account three types of bias present in crop correspondents' returns. Since the procedure Mr. Sanderson follows does not attempt to isolate the separate effects of the three biases, the latter

[2] Mr. Sanderson's footnote 9.

[3] The average yields per acre for the two bases are as follows:

Base	*Manitoba*	*Saskatchewan*	*Alberta*
1908–1930	16.8	16.3	17.9
1908–1940	16	16	18

[4] The published condition figures for 1931–1940 were based on the 1908–1930 average yield. While those for 1921–1930 were based on the average of the preceding 10 years, Mr. Sanderson doubts that any adjustment is warranted in comparing them with those issued on the 1908–1930 base.

may be considered here as one problem. His method is the one commonly employed by the Division of Agricultural Statistics in the United States Department of Agriculture in interpreting their condition figures, and consists of studying the regression of final yield on the monthly condition figures reported by correspondents over a period of several months preceding harvest.

The writer investigated this method of adjusting for correspondents' bias prior to embarking on the weather-yield analysis, but found the approach wanting in several respects as applied to the western Canadian situation. The reasons for its shortcomings in that particular situation were not hard to find, and were primarily concerned with the underlying hypotheses. In contrast with the winter wheat crop in the United States, the spring wheat crop in western Canada is grown within a very short season. In addition, the year-to-year yield-per-acre variability is considerably more than double that of the United States crop. While the weather determinants of yield for the United States winter wheat crop are spread over a nine-month period, the principal determinants of yield in western Canada occur within a two-month period. In the latter area, preseasonal and spring developments to May 31 play but a minor role in the final determination of yields. The crucial developments come in June and July. Even if May seeding has been done in a practically dry seed-bed, normal June rains can bring the plant into the shot-blade stage, and if July precipitation is good, a bumper yield can result. If, on the other hand, temperatures are extreme, the plant may burn or ripen prematurely; if July rainfall is light, reduced yields are the result.

Because of the brevity of the growing season, and the very wide amplitude in yields resulting from weather developments within that period, it follows that it is an invalid operation to employ the regression of yield on the reported May 31 condition as a means of isolating and adjusting for the bias in the correspondents' returns. The latter report on current condition, not on their forecasts of final yield. In the absence of any bias, one would expect at most to find but a very poor relation between final yield and "true" May 31 condition. The major determinants of yield have yet to occur. What Mr. Sanderson has isolated by his regression of yield on the reported May 31 condition is actually an inseparable admixture of correspondents' bias and the real discrepancies between final yield and "true" May 31 condition, of which the latter are of much the greater magnitude at that date.

A similar regression of yield on the June 30 reported condition is open to the same objection in principle, but modified by the fact that some of the more important determinants of yield are then on record. Even with normal June developments, however, July remains a crucial

month. Only a partial relation could be expected between yield and "true" *June 30 condition, and again the regression would isolate* an admixture of bias and real discrepancies.[5] At July 31, however, the major determinants are on record and it would then be a logical procedure to employ the regression of yield on condition as a means of adjusting for bias.

For the foregoing reasons the writer rejected the methods now proposed by Mr. Sanderson as a means of improving on the old-style condition figures for May 31 and for June 30. For the same reasons the writer is not inclined to attach significance to the correlation coefficients or indexes calculated by Mr. Sanderson for the correspondents' condition figures for those months. The only point remaining at issue is whether the July 31 condition figures afford, with proper adjustment, a better basis for approximating final yield than did the condition figures based on weather factors to which the change was actually made.

As evidence on this point, Mr. Sanderson makes a comparison for two of the three Prairie Provinces, on the basis of squared correlation indexes which he has calculated both for the July 31 condition based on correspondents' returns and for the July 31 condition based on weather factors. His comparison shows better results for the correspondents' returns in Saskatchewan and for the weather factors in Alberta. The squared correlation indexes he shows for both series, however, are open to some question.

In the first place, in developing the weather yield analysis, data for four years in Manitoba and one year each in Saskatchewan and Alberta were excluded because extraneous factors such as rust were dominant in those years. Their inclusion would have vitiated an approximation to the "true" net regressions of yield on the various weather factors. From that viewpoint, their exclusion was logical, but from Mr. Sanderson's viewpoint, it is necessary to compare the results of the weather analysis in *every* year of the 1921–1937 period for which he is showing his interpretation of the crop correspondents' returns.[6] The squared correlation indexes he calculated for the weather analysis implied that no adjustment whatever in practice would be made in years when extraneous factors become dominant. A more reasonable assumption would be that on the basis of the Bureau's numerous contacts with

[5] *If an approximation to correspondents' bias at May 31 and June 30 were desired one might well* study the differences between the correspondents' condition figures for these dates and the corresponding condition figures furnished by the weather-yield analysis. The latter are based on the partial regressions of yield on preseasonal and spring weather factors and, with allowance for their standard errors of estimate, are approximations to the "true" condition for these dates. The extent to which correlation is lacking between yield and "true" condition as of May 31 and June 30 emphasizes the logical futility of employing such data as a forecasting device. That the U.S.D.A. can use their data for forecasting purposes is undoubtedly due to the fact that their crop determinants are not nearly as heavily concentrated in the two months prior to harvest. The "true" condition figures under such circumstances would more closely approximate the final yield.

[6] See footnote 12 of Mr. Sanderson's article.

actual field developments, arbitrary adjustments would be made as close on the average as the standard error of estimate to the true adjustment for those exceptional occasions. From theoretical considerations it would follow that substitution of the standard error of estimate for the "nonsense" observations in the excluded years would leave the correlation indexes practically unchanged from those which were calculated after exclusion of the few years in which non-weather factors were dominant. On these grounds the writer rejects the July 31 correlation indexes for the weather analysis which Mr. Sanderson has calculated, and will use in comparison below the indexes as published.[7]

Secondly, just as Mr. Sanderson has asked that the years of "extraneous influence" be taken into account, he should be willing to consider the 1921–1940 period for which the data are available, rather than cutting off at 1937. The comparison shown here accordingly includes the correlation indexes calculated for the whole of the 1921–1940 period.[8]

Thirdly, one wonders why Mr. Sanderson has excluded the Manitoba case from consideration. It isn't sufficient to say that Manitoba accounts for only one-tenth of the Canadian wheat acreage. The problem of a proper basis for the publication of condition figures is just as relevant to the Manitoba situation as it is to that of the other two provinces. With Manitoba included, the comparison follows:

Period 1921–1940	Squared correlation indexes (corrected)		
	Manitoba	Saskatchewan	Alberta
July 31 condition based on correspondents' returns	0.363	0.767	0.908
July 31 condition based on weather factors	0.651	0.805	0.870

Surely there is no evidence here that the users of condition figures were misled by the adoption of the weather-yield approach, instead of adjusting for correspondents' bias in the old series in the way Mr. Sanderson has suggested.

(5) Mr. Sanderson's argument on this count is that if the Bureau had superimposed on the regression of yield on the old-style condition figures a further adjustment for certain weather data, as he has done in a forthcoming study,[9] results superior to those of the Bureau's weather analysis could have been obtained for forecasting purposes. The first objection to this procedure is the one already made against the use of regression of yield on condition at the May 31 and June 30 dates. The residuals against which the weather data would be plotted

[7] *Quarterly Bulletin of Agricultural Statistics*, King's Printer, Ottawa, Vol. 31, 1941, p. 187.

[8] In calculating the squared correlation indexes of the regressions of yield on correspondents' condition, the writer did not have access to Mr. Sanderson's regressions for the 1921–1937 period. The writer's curve for the 1921–1940 period, however, gave the same results for 1921–1937 as those shown by Mr. Sanderson.

[9] The writer greatly appreciates the effort made by the Division of Agricultural Statistics, U.S.D.A., to place in his hands a manuscript copy of Mr. Sanderson's study referred to in the latter's footnote 13,

are not a valid approximation to the correspondents' bias for which the weather data are employed to correct. Secondly, one wonders at the particular choice of weather factors Mr. Sanderson has employed. Preseasonal precipitation in conjunction with May 31 condition appears logical, in conjunction with June 30 condition subject to some question, and in conjunction with July 31 condition quite illogical since correspondents at that date would have visual evidence of any significant differences preseasonal precipitation has made. Ten days' preceding precipitation and temperature, if not highly intercorrelated for such a short period is a logical choice for June 30, but it would be just as logical for July 31. That the variables as they were employed have led to "statistically" significant improvements for the period under review holds questionable promise for their use as a forecasting device when the causal relations are not thoroughly established. Thirdly, the fact that higher correlation indexes have been adduced for June 30 than for July 31 in itself points to weakness in the underlying hypotheses.

(6) The inference conveyed by Mr. Sanderson's footnote 16 is misleading and unfortunate. In the study cited, the data were employed to establish none of the major hypotheses in the analyses for any of the three provinces, such as the partial regressions of yield on June and July precipitation and temperatures, nor in the case of April and May precipitation and temperatures in the case of Saskatchewan. When the analyses for Manitoba and Alberta were undertaken, inspection of data for these early factors was made in determining whether their inclusion was warranted. To this minor extent the data were used for both hypothesis and proof. Mr. Sanderson himself appears to have made similar use of data in choosing between linear and curvilinear regressions[10] and in some of the particular choices of weather variables employed in his forthcoming study, *à propos* the selection of which reference has already been made.

Mr. Sanderson's remarks about the potential pitfalls of multiple correlation analysis when several variables are used are of general application to the method. When these considerations are addressed to the particular application of the method in the Bureau's weather-yield analysis, it should be conceded that some of the common difficulties are avoided. The problem is one of physical rather than economic phenomena, and the causal relations are clear. Among the variables used, the intercorrelations are accidental and predominantly insignificant.

EDITOR'S NOTE.—It is anticipated that this discussion will be continued in the forthcoming March issue.

[10] See Mr. Sanderson's Table 11.

DISTRIBUTION OF INCOME IN 1935–36

By Rufus S. Tucker
General Motors Corporation

In two previous articles[1] the present writer has pointed out some deficiencies in the National Resources Committee's reports on the distribution of income in the United States in 1935–36. Every criticism made in those articles has been confirmed by subsequently published information, but the tentative revision suggested in the first article has turned out to be insufficiently drastic. Consequently the distribution here presented has been prepared, making use of the finally published results of the Consumers' Purchase Study, the detailed *Supplements to Statistics of Income for 1936*, and the reports of the Wisconsin and Delaware state income taxes for that year. Delaware and Wisconsin are the only states that have income tax statistics in usable form; one has a large proportion of large incomes, the other a small one.

Families. The NRC distribution for families over $7,500 was based on *Statistics of Income for 1935*, with certain assumptions as to evasion, understatement, capital gains and losses, the effect of separate reporting by married couples, and the shift in income between 1935 and 1936. The *Statistics of Income for 1936* and the special *Supplements* published for that year permit an accurate evaluation of separate reporting, and capital gains and losses, and the effect of adding back legally allowable deductions, and enable the analysis of income tax reports to be carried down to families with $3,500 income. This analysis can be readily converted to the 1935–36 period by comparison of the regular *Statistics of Income* for the two years.

If the NRC allowance for tax evasion and understatement was not excessive, and if the replacement of allowable deductions was justified, and the allowance for imputed income approximately correct, the NRC figure of 471,000 for the total number of families above $7,500 can be supported by the income tax data. But if the same method had been applied to all tax returns, instead of only to those with statutory net incomes of $5,000 or over, the number of families with economic incomes above $3,500 would have been 2,080,000 instead of the NRC figure of 1,635,000. Undoubtedly 2,080,000 is too large a figure, not being supported by any other evidence; whence it follows that 471,000 is probably too high an estimate for families above $7,500. Certainly there was no less tax evasion and probably no less understatement, in proportion

[1] *Review of Economic Statistics*, November 1940 (Vol. XXII, pp. 165–182) and February 1942 (Vol. XXIV, pp. 9–21).

to actual incomes, by taxpayers with statutory net incomes under $5,000 (or economic incomes under $7,500) than by those with larger incomes.

In each statutory net income class under $10,000 the assumption that 25 per cent of the citizens failed to file returns, and that the average returns filed understated net income by 15 per cent, though startling, might possibly be justified if taken by itself, as a measure to adjust statutory net income. But applied to total income, as it was by the NRC, it involves a double allowance for understatement. Perhaps the easiest and commonest way to understate taxable net income is to exaggerate the allowable deductions. By restoring these deductions the NRC automatically wiped out a large part of the understatement. In addition, by restoring all deductions the NRC refused to recognize that a large part of them were legitimate business expenses, which must be deducted to ascertain the correct amount of economic income. Consequently the number of incomes in the income classes above $7,500 was greatly exaggerated.

If the federal tax returns are adjusted for capital gains and tax-exempt interest and further adjusted by allowing for evasion and understatement on the NRC assumption, but without adding back statutory deductions, we get 380,000 for the number of families over $7,500 and 1,693,000 for the number over $3,500. These figures also include an allowance for imputed income, based on the average amount of imputed money value of current expenditures reported for each income class in Table 1 of *Family Expenditures in the United States.*

If no separate allowance is made for tax evasion and understatement, but total income is taken as the base instead of statutory net income, the number of families over $7,500 would be 300,000 and the number over $3,500 would be 1,130,000. These figures also include adjustments for capital gains and losses and tax-exempt income and non-money income.

The Delaware state tax returns, if adapted to the national population, would warrant taking 1,736,000 or more as the proper total for families over $3,500 without requiring any allowance for evasion or understatement, but Delaware has a greater proportion of large incomes than most states. The Wisconsin state tax returns would require a considerable allowance for evasion to come anywhere near that figure. Some such allowance must certainly be made, and it is almost as certain that the allowance should be greater for small incomes than for large ones. Consequently the figure of 380,000 has been preferred for $7,500 and over, and the figure of 1,711,000 for $3,500 and over.

This figure, 1,711,000, is based on the Consumers' Purchase Study,

TABLE I

DISTRIBUTION OF NON-RELIEF FAMILIES BY INCOME CLASSES 1935-36

Income class	Per cent As stated by NRC* (Weighted averages of localities) (233,000 cases)	Per cent As shown by CPS samples (Unweighted pool of all cases reported) (260,000 cases)
Under 250	2.83	2.10
250- 500	7.83	5.41
500- 750	11.30	8.59
750-1,000	13.40	11.33
1,000-1,250	13.10	12.77
1,250-1,500	10.80	10.08
1,500-1,750	9.05	9.08
1,750-2,000	7.34	9.15
2,000-2,250	5.46	6.07
2,250-2,500	4.01	5.20
2,500-3,000	5.23	9.06
3,000-3,500	2.09	3.90
3,500-4,000	1.76	2.21
4,000-4,500	1.00	1.55
4,500-5,000	.61	.77
5,000-7,500	1.30	1.00
7,500 and over	1.89†	.09

* Calculated from *Consumer Incomes*, p. 97, Table 9B.
† Adjusted by NRC on basis of income tax returns.

which was used by the NRC before the results of the survey were available to the general public.[2] The NRC used only part of the available material, and weighted it in accordance with various assumptions, some of which the present writer questioned in his previous articles. Table I has been prepared to show to what an extent the processing of the material by the NRC resulted in a different picture from that presented by the original data. The chief reasons for the differences were: (1) excessive weighting of the samples from the South, especially South Carolina and Georgia; (2) omission of some samples from more prosperous regions; (3) excessive weighting of broken families, with no distinction between those lacking a wife and those lacking a husband.

Both the unweighted CPS pattern and the NRC weighted pattern suffer from the non-representative nature of the communities and families sampled, although not to the same extent. Both also probably suffer from an inadequate allowance for non-money income of farmers and servants and from wide-spread understatement of income by the families interviewed, but no correction is possible on those accounts.

[2] The results of the Study of Consumer Purchases are now available in *Bureau of Labor Statistics Bulletin Nos.* 642-646 and *Department of Agriculture Miscellaneous Publications Nos.* 339, 345, 356, 370, 375, 384, 390, 399, 402, 405, 415, 422, 428, 432, 436, 452, 455, 456, 457, 462, 464, 465.

Since the published CPS samples refer only to non-relief families it is necessary to prorate them against the 24,913,000 non-relief families reported by the NRC, and to add to each class the NRC estimates of families on relief. The totals thus obtained are shown in Table II as cumulations for all classes from zero to $3,500. Above $3,500 the cumulations are based on the income-tax analysis.

TABLE II

DISTRIBUTION OF FAMILIES AND SINGLE INDIVIDUALS BY INCOME CLASSES IN 1935-36 (cumulated)

(in thousands of units)

Economic income	Families		Single individuals		Consumer units	
	NRC	Revised	NRC	Revised	NRC	Revised
Total	29,400	29,400	10,058	10,058	39,458	39,458
Over $500	25,222	25,997	7,535	7,760	32,757	33,757
" 750	21,423	22,876	5,553	5,926	26,976	28,802
" 1,000	17,149	19,114	3,951	4,314	21,100	23,428
" 1,250	13,263	15,328	2,845	2,860	16,108	18,188
" 1,500	10,398	12,417	1,967	1,967	12,365	14,384
" 1,750	8,051	9,843	1,424	1,432	9,475	11,275
" 2,000	6,157	7,529	1,022	1,091	7,179	8,620
" 2,250	4,737	5,700	738	749	5,475	6,449
" 2,500	3,693	4,368	528	539	4,221	4,907
" 3,000	2,378	2,698	367	361	2,745	3,059
" 3,500	1,635	1,711	258	252	1,893	1,963
" 4,000	1,149	1,217	191	188	1,340	1,405
" 4,500	942	1,006	159	158	1,101	1,164
" 5,000	794	817	133	133	927	950
" 7,500	471	350	76	65	547	415
" 10,000	284	228	47	41	331	269
" 15,000	152	111	26	22	178	133
" 20,000	93	68	17	14	110	82
Median	$1,160	$1,300	$815	$875	$1,070	$1,170
Lower tercile	$840	$970	$605	$615	$760	$870
Upper tercile	$1,670	$1,755	$1,120	$1,170	$1,450	$1,680
Highest decile	$2,750	$2,900	$2,010	$2,070	$2,600	$2,725
Highest percentile	$9,800	$8,650	$6,100	$5,600	$9,000	$8,100

As a check on the figures obtained by adding the CPS samples the Delaware and Wisconsin state tax returns for 1936 can be used. The federal tax returns are of no value for families below $3,500. The Wisconsin returns are also incomplete below $1,750, since returns were not required for statutory net incomes under $1,600.

These state returns, if adjusted by the methods used by the NRC for federal returns in the higher brackets would both show a greater proportion of families above $2,250 than the unweighted samples and a greater proportion above $1,750 than the NRC distribution. The NRC

allowance for evasion and understatement would probably be excessive for families in income classes over $2,000 in those states, where the presence of both federal and state collectors and auditors might be expected to reduce evasion to less than the national average. Likewise family incomes in 1939 were doubtless somewhat higher than in 1935–36. Nevertheless with very moderate allowances for evasion and no allowances for understatement the Wisconsin returns, although incomplete below $1,750, would indicate that in 1935–36 the number of families between $1,500 and $2,250 was greater than shown by using the samples, and the Delaware returns would indicate a greater number between $1,250 and $1,750, and between $2,000 and $2,250, and larger cumulative totals at every point above $1,250.

Taken together the state tax returns show plainly that the NRC estimate of families between $1,250 and $3,000 was too low and that the unweighted samples do not increase their number unduly.

Single persons. The distribution of single persons over $3,000 as reported by the NRC was based on federal income tax returns for 1935 with adjustments for families. The same objections therefore apply. In addition the allowance for imputed income was not stated in the NRC reports, because it was "not considered sufficiently reliable to warrant presentation."[3]

The revised figures used herewith for classes over $10,000 are based on the 1936 *Supplements*, adjusted to the average of 1935 and 1936 by the ratio of the number of returns in each class as reported in *Statistics of Income* for those years. They represent the higher of two figures, one of which is based on total income adjusted for capital gains and losses and tax-exempt interest; the other being statutory net income adjusted for evasion and understatement in accordance with the NRC formula, with tax-exempt interest added. Since evasion was doubtless greater among individuals than among families the number obtained by this method between $5,000 and $10,000 has been raised to conform to the respective number of families and individuals in that class reported in the *Supplement*.

For classes between $1,500 and $5,000 the NRC figures have been retained, with a little smoothing to make them compare better to the number of families in each class. Between $1,500 and $3,000 the NRC figures should perhaps be raised. They were almost entirely derived from samples relating to women only, and the NRC's adjustment to cover men was entirely a matter of judgment. Moreover the samples were mainly collected in the years 1931–34, the bottom years of the

[3] *Consumer Expenditures*, p. 147.

depression, and no adjustment appears to have been made to bring them into line with the higher wages and greater employment of 1935–36.

Between $1,000 and $1,500, the NRC figures must certainly be raised. They are too low by comparison with either the Delaware or Wisconsin tax returns if those returns are adjusted in the same way that the NRC adjusted the federal tax returns in the higher brackets. The numbers in each class between $750 and $1,500 were too low according to both states' returns if the number reported in each class is multiplied by a factor that makes the total above $1,500 equal to 1,067,000, the NRC figure. Finally the numbers between $1,000 and $1,500 were too low to be consistent with the number of families in those income classes, as revised. Obviously if the total number of single individuals is one-third of the number of families, the proportion must be greater than one-third in the lower income classes, and the proportions between the cumulated totals must approach one-third as the point of cumulation approaches zero. The federal and state tax returns, and the NRC published estimates, all agree that the number of single individuals increases faster than the number of families as one goes down from $2,500 or $2,000 to zero incomes. Taking the ratios of families to individuals shown by the NRC estimates, which are higher than those shown by other sources and therefore give a minimum number of single individuals, the increases in the aggregates of families over $500, $750, $1,000 and $1,250 already demonstrated require that the number of individuals in the corresponding aggregates be increased as shown in Table II. This results in increasing the number in each class between $1,000 and $1,500 and decreasing the number in each class below $1,000. If the state tax returns alone had been used as basis for revision, with very moderate adjustments for evasion the class $750–$1,000 would also have been increased and the cumulative total over $750 would have been larger than shown in Table II.

Classes below $500. No attempt is here made to subdivide the families or single individuals below $500. The CPS samples in this area are open to grave suspicion. In the Consumer Expenditures volume[4] the NRC rejected flatly the statements of non-relief families and single individuals in this class concerning their expenditures. If their reported expenditures could not be relied upon, were their reported incomes reliable? Moreover it is of little importance whether the incomes of non-relief families and single individuals under $500 were above or below $250 since the overwhelming majority of non-relief families and single persons in that class in any given year are there only temporarily, and in many cases in a technical sense only, because their business losses

[4] *Consumer Expenditures*, pp. 134, 140.

and deductible expenses happen to be abnormally large. They live on their capital or on borrowed money until their income picks up, and if it does not pick up they go on relief. If, as the NRC seemed to imply in its expenditure study, a family cannot live on less than $500 nor an individual on less than $250, using the NRC definition of income, there can obviously be no families or individuals with incomes less than $500 or $250, unless they have liquid assets or good credit. The NRC definition of income included gifts and inherited cash if spent for current living expenses; also non-money income from owned homes and home-grown food and fuel, and payments in kind, and public relief in cash or in kind. And since relief is not ordinarily granted to persons with enough liquid assets or good credit to live on, and is always granted to such an extent as to prevent starvation and nakedness, it would follow that most of the 1,523,000 relief families with incomes under $500 and the 329,000 single men and 169,000 single women on relief with incomes under $250 reported by the NRC either did not exist or else should have been placed in some higher class.

From the foregoing it appears that the total of 3,403,000 families and 2,298,000 single individuals left in the under-$500 class is probably too great, but there is no substantial basis of fact to warrant any specific lower figure. It is at any rate apparent that the NRC put at least 1,000,000 and possibly 2,500,000 too many consumer units into the class with incomes under $500.

The pattern of distribution of families published by the NRC, although purporting to be based on the Study of Consumer Purchases and the income-tax statistics, differed widely from its sources and reflected more the result of various assumptions and conjectures. It greatly exaggerated the numbers over $7,500 and under $1,250. The tabulation here presented conforms more closely to the original data, but may still exaggerate the numbers over $7,500 and under $1,250 and especially under $500.

The NRC pattern of distribution of single individuals in 1935-36 was based on inadequate data, largely collected in years less prosperous than 1935-36. Checked against federal and state income tax data it appears to have overstated the numbers over $7,500 and below $750. The present tabulation perhaps still does so, but to a lesser extent.

The NRC pattern of consumer units is so far removed from the original data and from the probable facts that it is highly unlikely that any conclusions drawn from it, unless adequately supported by other evidence, can be sound. For the sake of scientific truth and sound public policy, its further use by economists should be avoided, and the many studies already published relying on it should be revised.

A PUNCHED CARD TECHNIQUE TO OBTAIN COEFFICIENTS OF ORTHOGONAL POLYNOMIALS

By Lila F. Knudsen[1]
Food and Drug Administration

In past years many methods have been developed for fitting polynomials. Although the use of orthogonal polynomials for fitting continuous functions probably originated with Legendre, Tchebycheff[2] seems to be the first to use orthogonal polynomials for fitting discrete observations with and without equal intervals. Orthogonal polynomials have been the subject of many mathematical dissertations and an entire volume[3] of references and cross references on this subject has been compiled by a committee, of which J. Shohat is chairman. The fitting of polynomials by the use of orthogonal functions has been mainly by two methods: (1) summation; and, (2) multiplication of the variates by actual values of the orthogonal polynomials as given in a table. Pearson[4] and Isserlis[5] dealt with the fitting of non-equidistant and unequally weighted data. R. A. Fisher developed methods of calculation of Tchebycheff polynomials by successive summation up to the 5th degree,[6] and in his classic work on "Yield of Wheat at Rothamsted"[7] refers to work by Esscher.[8] F. E. Allan published a paper, "The General Form of Orthogonal Polynomials for Simple Series."[9] C. Jordan[10] develops the theory in full, and crediting the method to Tchebycheff in a paper published in Russian in 1926, demonstrates the building up of general expressions for the orthogonal polynomials in factorial form originally used by G. F. Hardy in "Graduation of British Offices Tables, 1863–93." Jordan also suggests the numerical method of building up the polynomials by summation and furnishes tables for this purpose. J. Shohat gave the mathematical approach in "Stieltje's Integrals in Mathematical Statistics."[11] A. C. Aitken[12] gives a relatively simple demonstration of the theory and includes a series of tables up

[1] The author wishes to thank Dr. Harold Hotelling of Columbia University for his helpful suggestions and criticisms, and Mr. Meyer A. Girshick of the Bureau of Agricultural Economics in the Department of Agriculture for his continued encouragement, suggestions and assistance.

[2] P. L. Tchebycheff, 1854–1875, see *Oeuvres* (1899).

[3] J. Shohat (Chairman) "Bibliography on Orthogonal Polynomials," assembled by Committee of National Research Council, *Bulletin of the National Research Council*, No. 103, August, 1940.

[4] K. Pearson, *Biometrika*, Vol. 13, pp. 296–300 (1921).

[5] L. Isserlis, *Biometrika*, Vol. 19, pp. 87–93 (1927).

[6] R. A. Fisher, *Journal of Agricultural Science*, Vol. 11, pp. 107–135 (1921).

[7] R. A. Fisher, *Philosophical Transactions of the Royal Society*, B 213, pp. 89–142 (1924).

[8] F. Esscher, "Über die Sterblichkeit in Schweden (1886–1914)," "Meddelanden från Lunds Astronomiska Observatorium," 2, 23, pp. 10–21 (1920).

[9] F. E. Allan, *Proceedings of the Royal Society of Edinburgh*, Vol. 50, p. 310 (1929–30).

[10] C. Jordan, *Annals of Mathematical Statistics*, Vol. 3, pp. 257–357 (1932).

[11] J. Shohat, *Annals of Mathematical Statistics*, Vol. 1, pp. 73–94 (1930).

[12] A. C. Aitken, *Proceedings of the Royal Society of Edinburgh*, Vol. 53, pp. 54–78 (1933).

to the 5th degree for the summation method. Aitken, Fisher, and F. F. Stephan[13] show individually the advantage of dividing a series of data and summing from the extremes. Two excellent resumés of the subject of orthogonal polynomials with many references are given by Wishart[14] and Irwin.[15]

Usually the objection raised to fitting orthogonal polynomials to a large mass of data is the laboriousness of the task. This objection may be overcome by using punch cards and Hollerith machines, namely, card punch, sorter, tabulator and summary punch. This article will demonstrate fitting of orthogonal polynomials to weekly weighings of rats by means of tabulating equipment, and will discuss the analysis of the coefficients of the polynomials in order to investigate the effects of animal diets.

A process of fitting orthogonal polynomials by a punched card method of successive summations has been published by Richard Warren of Columbia University.[16] However, his method involves at least as many sets of cards and as many runs as the degree of the polynomial. For instance, for fitting a polynomial of the sixth degree, six sets of cards would be used, and six runs besides the three or four runs for calculating sums of squares. This method seems cumbersome and lengthy for fitting polynomials simultaneously to more than one series of data.

The method described in this article uses the direct method of multiplying by the tabled actual values of the orthogonal polynomials and is relatively short involving only about one-fifth the time and work required by Mr. Warren's method. The author is not aware of any previous description of fitting orthogonal polynomials by the direct method of multiplication by addition on a punched card accounting machine using the digiting process.

The procedure is given here not only because of its use in fitting orthogonal polynomials but also because of its adaptability to other problems wherein values obtained from tables are multiplied by one or more variates in order to obtain some type of curve. The method has recently been used by Miss Dorothy Nickerson, Cotton Branch, Agricultural Marketing Administration, to obtain tristimulus values for 425 Munsell colors for each of four illuminants using the method described in Hardy's *Handbook*.[17] The entire process of punching, tabula-

[13] F. F. Stephan, this Journal, Vol. 27, pp. 413–423 (1932).

[14] J. Wishart, *Journal of the Royal Statistical Society*, Vol. 96, pp. 487–491 (1933).

[15] J. O. Irwin, *Journal of the Royal Statistical Society*, Vol. 97, p. 134 (1934).

[16] R. Warren, "Hollerith Machine Computation of Least Square Trend Lines," *Columbia University Statistical Bureau Document* #2 (1933).

[17] A. C. Hardy, *Handbook of Colorimetry*, Massachusetts Institute of Technology (1936).

tion and calculation of the tristimulus values for each of the four illuminants from the 425 spectrophotometric curves (one for each of the Munsell colors), was completed in less than one month, at about one-eighth of the cost estimated to do the job with a multiplying punch, whereas two or three years would have been required to complete the work using ordinary methods of calculation.

Various methods have been used to compare effects of different diets on animals. Some research workers have compared average growth of experimental and control animals over a definite period. Others have merely visually compared growth curves, etc. A more precise method of making use of all the weekly weighings is to use some algebraic expression for the growth curve, and compare the growth of the animals on various diets by comparing the coefficients of their growth equations. Various types of equations were tried by the author on this type of data but orthogonal polynomials seemed to fit the actual data better, and moreover one can add or delete a higher power without recalculation of the coefficients of preceding powers. While this study was being developed in the Food and Drug Administration, Wishart published a paper[18] wherein he fitted orthogonal polynomials to the individual weekly weights of pigs over a period of 17 weeks. He studied the effects of the diets by applying the analysis of variance technique to the orthogonal polynomial coefficients, a technique similar to that used on coefficients of Fourier series fitted to economic data.

Instead of fitting the polynomial

$$y = a + bx + cx^2 + dx^3 + \cdots$$

to a series of observations at equal intervals (where y is the weight of the rat at any week number in our present problem and x is the week number) Fisher[19] proposed a modification of Tchebycheff's method by using an equivalent polynomial

$$Y = A + B\xi_1 + C\xi_2 + D\xi_3 + \cdots$$

where $\xi_1, \xi_2, \xi_3, \cdots$ are orthogonal polynomials of degrees 1, 2, 3, $\cdots$ and where $A = \bar{y}$ and $\xi_1 = x - \bar{x}$, $\xi_2 = a'x^2 + b'x + c'$, etc.

The coefficients of the orthogonal polynomials are found by calculating the sums of the products of the y values and the ξ' values (where y equals the observed weight for any one week) and dividing by the sum of squares of the ξ' values (which are given in Fisher and Yates tables[20]).

[18] J. Wishart, "Statistical Treatment of Animal Experiments," *Supplement to the Journal of the Royal Statistical Society*, Vol. VI, No. 1, 1939.

[19] R. A. Fisher, *Statistical Methods for Research Workers*, Oliver & Boyd, Ed. 6, p. 139 (1936).

[20] Table XXIII of *Statistical Tables for Biological, Agricultural, and Medical Research*, by R. A. Fisher, and F. Yates, Oliver & Boyd (1938), gives numerical values for ξ' (where $\xi' = \lambda\xi$) for all ξ' from

The coefficients are given by the algebraic expressions:

$$B' = \frac{S(y\xi_1')}{S(\xi_1'^2)} \qquad C' = \frac{S(y\xi_2')}{S(\xi_2'^2)} \quad \text{etc.}$$

Since the ξ''s are merely multiples of the ξ values, the $A, B, C, \cdots$ may be easily obtained from the $A', B', C', \cdots$ values.

In order to adapt this method to Hollerith machine punch cards, it would be necessary first to code the values given in the Fisher and Yates tables by the addition of some power of ten so as to avoid negative numbers. This can be illustrated by giving part of the table of ξ' values for $n'=52$, as given in the Fisher and Yates tables, page 60, together with the ξ'' which are the coded values of the ξ' (see Tables I and II). The values of A, B, C, etc., can be obtained by using the relationships

$$A = \frac{S(y)}{52}$$

$$B\xi_1 = \frac{S(\xi_1'Y)}{46{,}852}\cdot\xi_1' = \frac{2S(\xi_1'y)}{46{,}852}\cdot\xi_1 = \frac{S(y\xi_1'') - 100Sy}{23{,}426}\,\xi_1$$

$$C\xi_2 = \frac{S(y\xi_2'') - 1000Sy}{2{,}108{,}340}\,\xi_2$$

$$D\xi_3 = \frac{S(y\xi_3'') - 10{,}000Sy}{243{,}513{,}270}\,\xi_3$$

$$E\xi_4 = \frac{S(y\xi_4'') - 10{,}000Sy}{2{,}597{,}474{,}880}\,\xi_4$$

$$F\xi_6 = \frac{S(y\xi_6'') - 100{,}000Sy}{527{,}169{,}333{,}600}\,\xi_6.$$

The quantity $S(y\xi''_{1,2,\cdots,6})$ may also be calculated and the identity $S(y\xi''_{1,2,\cdots,6}) = S(\xi_1''y) + \cdots + S(\xi_6''y)$ used to check the results. As a further check, the relationship

$$Z = B + 90C + 10{,}395D + 110{,}880E + 22{,}503{,}600F$$

may be employed where

$$Z = \frac{S(y\xi''_{1,2,\cdots,6}) - 121{,}100Sy}{23{,}426}.$$

ξ_1' to ξ_5', and for all n' from 3 to 52. R. L. Anderson and E. E. Houseman of Iowa State College at Ames, Iowa, have recently published an auxiliary table of ξ' values for larger values of n', "Tables of Orthogonal Polynomial Values extended to $N=104$," Iowa State Agricultural Experiment Station Statistical Section, *Research Bulletin* No. 297.

TABLE I

TABLE OF ξ' VALUES AS GIVEN BY FISHER AND YATES

Obs. No.	ξ_1'	ξ_2'	ξ_3'	ξ_4'	ξ_5'
1	−51	425	−4 165	3 570	−55 930
2	−49	375	−3 185	2 170	−23 030
3	−47	327	−2 303	1 022	658
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
51	49	375	3 185	2 170	23 030
52	51	425	4 165	3 570	55 930

TABLE II

TABLE OF CODED ξ' VALUES OR ξ'' VALUES

Obs. No.	ξ_1''	ξ_2''	ξ_3''	ξ_4''	ξ_5''	$\xi''_{1+2+\cdots+5}$*
01	049	1 425	05 835	13 570	044 070	064 949
02	051	1 375	06 815	12 170	076 970	097 381
03	053	1 327	07 697	11 022	100 658	120 757
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
51	149	1 375	13 185	12 170	123 030	149 909
52	151	1 425	14 165	13 570	155 930	185 241

* Check column $\xi''_{1+2+\cdots+5} = \xi_1'' + \xi_2'' + \cdots + \xi_5''$.

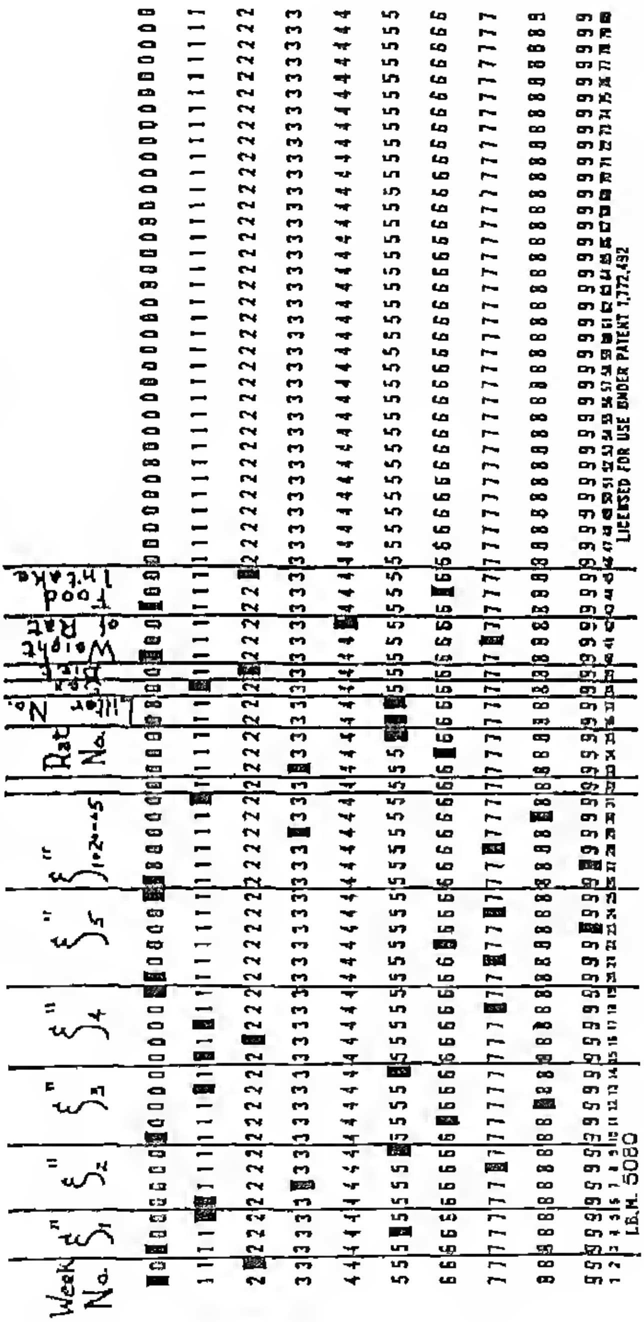
The use of punch cards facilitates the calculation of the Sy, $S(y\xi_1'')$, $S(y\xi_2'') \cdots$. One set of 52 cards is punched with the observation (or week) number, the ξ_1'', ξ_2'', $\xi_3'' \cdots$ values, and the check value $\xi_{1+2+\cdots+5}$. Referring again to Table II, one card is made for each *row* (not column). By means of a gang punch as many sets (of 52 cards each) are duplicated as there are curves to be fitted. Then taking one set of cards, it is a simple matter to punch in each card the rat or animal identification number, the litter number, the weight in grams of the animal, and its food consumption in grams for that particular week.

In order to illustrate more clearly the method used, an example will be given using data from an experiment conducted in the Food and Drug Administration.[21] In this experiment four different diets were fed to rats for the purpose of ascertaining the effect of the diet on the growth of the rat. Four rats were selected from a litter at weaning and one rat placed on each of the diets at random. Individual ad lib feedings were used. Chart I shows a card punched for week number 2 for rat number 365. (Incidentally, printed card forms are not necessary, plain cards may be used and a punched card code kept on file.)

[21] Data used in the description of this method were obtained from the Division of Pharmacology (Dr. Herbert O. Calvery, Chief) in the Food and Drug Administration where a large number of chronic toxicity studies are being made over long periods of time. The details of the experimental findings will be published elsewhere.

CHART 1

PUNCHED CARD FOR WEEK NUMBER 2 OF RAT NUMBER 365



The sums of cross products can be obtained mechanically in the following manner. The punch cards for all rats are sorted (using the Hollerith sorter) on the units column of the weight and removed from the sorter in reverse order (i.e., the 9's first, then the 8's, etc.), then sorted on the rat number. Having the tabulator set for progressive totals, tabulations are made on each of the ξ'' values, as illustrated for rat number 305 by the first part of Table III, clearing the dials of the

TABLE III

Wt. of Rat Units Col.	Rat No.	Card Count	S_1	S_2	S_3	S_4	S_5	[illegible]
9	305	5*	000113	0005781	[illegible]	[illegible]	[illegible]	[illegible]
8	305	10*	000712	0010678	[illegible]	[illegible]	[illegible]	001175830
7	305	17*	001423	0016215	[illegible]	[illegible]	[illegible]	002051019
6	305	22*	001984	0022230	[illegible]	[illegible]	[illegible]	002657130
5	305	26*	002362	[illegible]	[illegible]	[illegible]	[illegible]	003190709
4	305	33*	003155	[illegible]	0031247	[illegible]	[illegible]	003885721
3	305	39*	003611	[illegible]	[illegible]	[illegible]	[illegible]	004635549
2	305	46*	004021	[illegible]	[illegible]	[illegible]	[illegible]	005585012
1	305	48*	004744	0018522	[illegible]	[illegible]	[illegible]	005835998
0	305	52	005200	0052000	00520000	[illegible]	[illegible]	006207200
Tens Col.								
90	305	2*	001610	0017120	[illegible]	[illegible]	[illegible]	001973140
80	305	4*	003000	0035110	[illegible]	[illegible]	[illegible]	001510550
70	305	7*	001900	0070720	[illegible]	[illegible]	[illegible]	007829510
60	305	10*	006960	0103140	[illegible]	[illegible]	[illegible]	011560820
50	305	12*	008160	0126120	[illegible]	[illegible]	[illegible]	011012100
40	305	14*	009320	0151490	[illegible]	01375410	013740230	016709520
30	305	18*	014120	[illegible]	01801910	[illegible]	[illegible]	021374120
20	305	32*	031720	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
10	305	43*	013190	0141870	[illegible]	[illegible]	[illegible]	052851030
00	305	52	052000	0520000	05200000	05200000	[illegible]	[illegible]
Hundreds Col.								
300	305	35*	409100	3325300	[illegible]	[illegible]	[illegible]	420550100
200	305	46*	487000	4116000	[illegible]	[illegible]	[illegible]	556072000
100	305	50*	510000	4920000	[illegible]	[illegible]	[illegible]	613487000
000	305	52	520000	5200000	[illegible]	[illegible]	[illegible]	620720000

Totals of Summary Cards

305		1653282	14388208	[illegible]	[illegible]	[illegible]	1701011020

Orthogonal Polynomial Coefficients for Rat 305

$B = 3.2734$ $C = -.18058$ $D = .006590$ $E = -.0003161$ $F = .00000581$ $S(y) = 11760$ and $A = 283.062$

NOTE: Newer machines equipped with list rolling devices can produce the results shown without the use of summary cards, still further speeding up the process.

* Indicates summary cards used in the totals below.

tabulator after each rat number. By means of a summary punch, the rat number, the units column of the weight, and the progressive totals of the ξ'' columns are punched on summary cards as they are tabulated, one summary card for each row in Table III. The summary cards for the zeros on the units column of the weight are sorted out after comparing to see that they are identical for all rats—the figure in each column

on these "zero cards" should be 52 times the coding factor (which is in this case some power of 10). These punch cards are then discarded. These "zero cards" were checked and sorted out in the next two sets of progressive totals also. As indicated in Table III only the starred values are used in the totals at the bottom of the table. Caution must be exercised to insure a complete series of numbers for the progressive digiting, though the series may begin at, say, 3 or 300 instead of 9 or 900. If some number is missing in the progression, an extra *summary* card should be punched which is identical with the summary card preceding the omission and this extra summary card included with the other summary punch cards used to tabulate the totals of the summary cards.

The original punch cards are again sorted on the *tens* column of the weight and then on the rat number and tabulated the same as for the units column sort. Summary cards are also punched as before, however changing the wiring in the plug board so that a zero is automatically added after each total as shown in the second part of Table III headed "*tens column.*" The original punch cards are also sorted on the hundreds column of the weight and the same procedure followed as for the previous groups, this time two zeros are added automatically after each total on the summary cards as shown in the third section of Table III headed "hundreds column."

All the summary cards, not including the 3 groups of "zero cards" that were previously sorted out and checked, are then sorted on rat number and tabulated controlling on the rat number to obtain the totals of the summary cards for each rat. This gives the $S(y\xi_1'')$, $S(y\xi_2'')$, $S(y\xi_3'')\cdots$ which were used in the formulae for obtaining the B, C, D, $\cdots$ values. The entire process is illustrated in Table III. Summary cards used in the total are indicated in Table III by asterisks. The last line of the table indicates totals of summary punch cards or the $S(y\xi_i'')$. The original punch cards are also tabulated controlling on rat number to obtain total weight (Sy) and total food consumption. The method of progressive digiting which is a mechanical summation to obtain sums of squares and cross products is explained by Dr. Brandt.[22] Time can be saved using this punched card method even if only 10 or 15 curves are to be fitted.

As Wishart[23] points out, there is a biological interpretation to the coefficients of the orthogonal polynomials. The "A" coefficient for any one curve is the average weight of the rat over the 52 week period, the "B" coefficient is the average gain in weight per week, the "C" coeffi-

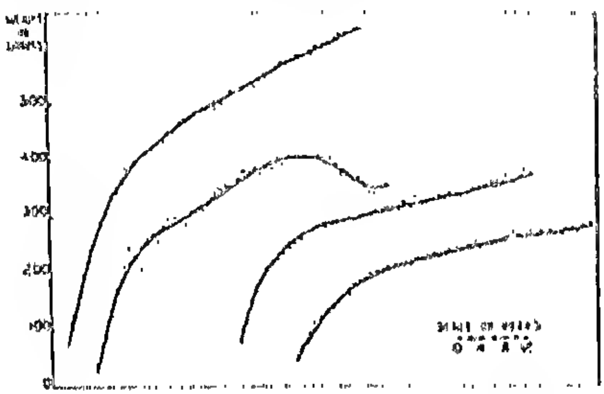
[22] *Practical Applications of the Punched Card Method in Colleges and Universities*, edited by G. W. Baehne, Columbia University Press (1935), Part X, pp. 423–436, entitled "Uses of the Progressive Digit Method," by A. E. Brandt.

[23] See footnote 18.

cient is proportional to the change in gain per week, etc. R. A. Fisher states in discussing Wishart's paper that the curve obtained using the averages of the coefficients of the orthogonal polynomials fitted to the individual animals is in fact identical with that obtained by a least squares fit to the weekly averages. The logarithms of the weights may

CHART II

FIFTH DEGREE POLYNOMIALS FITTED TO WEEKLY WEIGHINGS FOR ONE YEAR OF FOUR RATS SHOWING DIFFERENT DEGREES OF GROWTH VARIABILITY



be used instead of the weights; in fact, Wishart seems to prefer logarithms for the data used in his study.

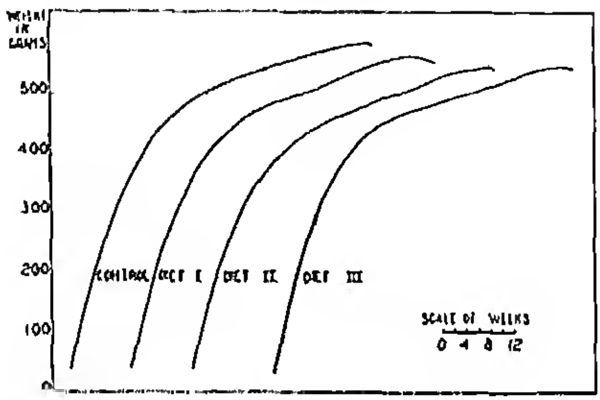
The question was raised as to the validity of the usual significance tests because of the interdependence of the observations. Stated another way, it was questioned whether the highly interdependent weekly weighings on one rat should be used to determine the degree of the curve fitted since the usual significance tests are based on an assumption of independence of the data. (Other aspects of this same question of serial correlation have been discussed by many writers. A partial list of references was given in an article in this JOURNAL[21] in 1940.) It may be argued that the interdependence is of a negligible amount in growth curves of rats. To investigate this question individual weighings on 100 comparable rats fed the same control diet over a period of one year were assembled. Using this data, instead of calculating for each rat the serial correlation coefficient as it is usually defined (i.e., the correlation between successive observations on the same animal) a slightly different type of correlation was employed, namely, the correlation between the weight of a rat one week and its weight some one following week, with

[21] L. F. Knudsen, *Interdependence in a Series*, this JOURNAL, Vol. 35, p. 507 (1940).

$N=100$. In this way a more definite picture could be obtained of the estimated amount of interdependence at various points on a general growth curve, after the average "trend" has been removed. These correlation coefficients varied from .47 between the weight of the rat the 3rd week and its weight the 26th week to .998 between the weights for

CHART III

AVERAGE GROWTH CURVE FOR EIGHT MALE RATS ON EACH DIET OVER A PERIOD OF 52 WEEKS



the 25th and 26th weeks. In no case was the correlation between adjacent weeks less than .9.

In the study of the Food and Drug Administration data which has been used as an example, fifth degree polynomials seemed to fit the data better than a polynomial of lower degree and as well as a polynomial of higher degree. Chart II shows the weekly weights of four rats for one year and the fifth degree polynomials fitted to them. The second and third curves in Chart II show the type of difference that may exist in growths of animals, while the overall gain in weight over that period may be identical. Gain in weight (the same as a first difference) assumes a straight line regression of weight with time. It can be seen from inspection of Chart II that *a considerable amount of serial correlation exists even in the residuals* from this curve, i.e., an observation above the curve is *usually followed by one or more successive observations above the curve*. The significance of the coefficients might be tested by the usual method using an error based on the sum of squares of residuals from a single curve. However, because of the high degree of interdependence between the residuals from the curve, it seemed desirable to use a test of the average coefficient based on the scatter of the coefficients of

several rats of the same diet and sex about that average. The averages and standard deviations of the individual coefficients (the A's, the B's, the C's, etc., separately) were calculated for each diet and sex. The standard error of the average A, B, or C, etc., coefficient was calculated in the usual way, that is, by dividing the standard deviation by the square root of N. In other words, the standard error of each average coefficient was calculated by considering the scatter of the *individual coefficients about the average*. As a test of significance of the coefficients, Student's "t" test was applied to test whether the average coefficient differed significantly from zero. This procedure avoids any assumption of independence of the weekly weights. Furthermore, the A, B, C, $\cdots$ coefficients are independently distributed.

Table IV gives an idea of the size of the average coefficients in the example used and their standard errors; and Chart III shows the growth curves described by these coefficients. The average coefficients includ-

TABLE IV

AVERAGE ORTHOGONAL POLYNOMIAL COEFFICIENTS FOR THE CONTROL AND THE THREE EXPERIMENTAL DIETS

Male Rats	A av $\pm\ \sigma_{av}$*	B av $\pm\ \sigma_{av}$*	C av $\pm\ \sigma_{av}$*	D av $\pm\ \sigma_{av}$*	E av $\pm\ \sigma_{av}$*
Control	461.77 ± 21.38	8.211 ± .509	−.2995 ± .0309	[illegible]	−.000273 ± .000037
Diet I	448.22 ± 20.58	7.673 ± .479	−.2784 ± .0214	[illegible]	.000173 ± .000021
Diet II	422.51 ± 25.10	7.471 ± .609	−.2812 ± .0111	[illegible]	−.000240 ± .000020
Diet III	438.07 ± 19.74	7.312 ± .510	−.2649 ± .0163	.01309 ± .00085	−.000301 ± .000040

* The average of the indicated coefficient for all the rats of the same sex on each diet ± the standard error of the average.

NOTE: This table is given merely for the purpose of illustrating the typical results to be expected and the size of the errors.

ing those of fourth and fifth degree were significantly different from zero, showing that a polynomial of not less than the fifth degree should be used in this problem. As contrasted with Wishart's study, where he uses a third degree polynomial for a 17-week growth period, it seems logical that a 52-week growth period would require a slightly more complicated curve.

The effects of the diets on the growth of the rat can be tested by using analysis of variance either on the coefficients themselves or on the coefficients corrected for the influence of initial weight and/or the food intake.

In growth studies there is good reason to expect a high degree of interdependence between residuals from growth curves. Consequently, instead of employing the usual tests of significance it may be advisable to estimate the standard error of the coefficients from the scatter of the individual coefficients about their average.

A NEW MULTIPLICATIVE SEASONAL INDEX*

By Julius Shiskin
National Bureau of Economic Research

This paper proposes a new multiplicative seasonal index, describes a quick way to compute it, and compares the new index with the ratio-to-moving-average index and with the monthly means index. The new index has some merits of its own, but the chief value of this paper lies in the light it throws on the relations among these indexes.

Derivation and computation. Let us consider a monthly time series for which a scatter diagram is drawn so that values for a given month (e.g., the January values) are plotted as the ordinate and the corresponding values representing the trend and cyclical components (hereafter referred to as the trend-cycle component) as the abscissa. If the original values for the month include neither a random nor a seasonal component, all the points fall on a straight line that passes through the origin and has a slope of one, since the trend-cycle component has merely been plotted against itself along axes with identical scales. If the assumptions are changed to allow a multiplicative seasonal component in the original values, all the points fall on a straight line that passes through the origin, but the slope deviates from one. If the original values include an additive seasonal component, the slope of the line remains one, but the line no longer passes through the origin. If the seasonal component is partly additive and partly multiplicative, the line does not pass through the origin and its slope differs from one. These relations tend to prevail if the series also includes a random component. However, the observations no longer fall on a straight line, but tend to be distributed at random around such a line. It can be concluded, therefore, that the seasonal component for a given month can be measured by the difference between the parameters of a fitted straight line and the parameters of a line passing through the origin and having a slope of one.[1]

The relations between the seasonal and the trend-cycle components may be expressed as follows. Start with the general linear equation

* This paper summarizes a section of a study on the measurement and elimination of seasonal fluctuations, in which particular attention is paid to the general method described below.

The writer is under heavy obligation to Dr. Arthur F. Burns for many contributions to this work. Valuable suggestions were also made by Dr. Simon Kuznets, Mr. Bert Gottfried, and Mr. Denis Volkenau.

[1] This approach to the study of seasonal fluctuations was suggested to the writer by Mr. Milton Friedman and Mr. Gerhard Bry. It appears in "Ausschaltung von Saisonschwankungen mittels Lag-Correlation," by F. L. Zassvy, *Monatsberichte des Oesterreichischen Instituts für Konjunkturforschung*, Beilage n. 2, Vienna, October 1933; and "Note on Seasonal Variation" by J. Wisniowski, *Studia Ekonomiczne*, Vol. 3, Cracow, 1936. It is a simplification of a method described by Horst Mendershausen in "Eliminating Changing Seasonals by Multiple Regression Analysis," *Review of Economic Statistics* November 1939, pp. 171–177.

$y_r = a + bx$, where y_r denotes a regression value for a given month as distinguished from y, an actual value of the original series, and x denotes the corresponding trend-cycle value. As long as there is no seasonal component, $y_r = x$. If the seasonal component is multiplicative, $y_r = bx$, and $b \neq 1$. If the seasonal component[2] is additive, $y_r = a + x$.

It is instructive to examine a special case of the general linear equation, the case where $b = 1$, or is assumed equal to 1. The value of a can now be evaluated for an empirical series from the equation:

$$\sum y_r = \sum y = na + b\sum x,$$

where n is the number of years covered by the given month. Here

$$a = \frac{\sum y - \sum x}{n} = \frac{\sum (y - x)}{n}.$$

The last expression is precisely the formula used for deriving an additive seasonal index by averaging actual differences of original values from the trend-cycle curve. Hence, in this special case, a is the uncentered[3] additive seasonal index.

Another special case is where a is zero, or is assumed equal to zero. Here $\sum y = b\sum x$ and $b = \sum y / \sum x$.

$\sum y / \sum x$ is suggested as a new multiplicative seasonal index.[4] It may be calculated as follows. (1) The sums of the original figures and the sums of the trend-cycle values for each month are computed. In the computations it is convenient to represent the trend-cycle component by the twelve-month moving average of the original series.[5] (2) The

[1] It is assumed here that the random factor has no disturbing effect upon these relations. This assumption is valid in principle, but not in a particular series.

[2] That is, the indexes for the twelve months (the discussion in the text refers to any one month) will not sum to zero. The indexes are centered by subtracting from each one-twelfth the difference between the sum of the uncentered indexes and zero. One consequence of this adjustment is that for any year the sum of the original figures equals the sum of the seasonally adjusted figures.

[4] It is important to note that the multiplicative seasonal factor is assumed here to apply to the trend-cycle component, not to the trend-cycle-random component as in the standard methods.

A simple procedure for converting the additive into the multiplicative index, and vice versa, is indicated by the expression obtained when the equation, $\sum y = na + b\sum x$, is divided through by $\sum x$:

$$\frac{\sum y}{\sum x} = a\frac{n}{\sum x} + b = \frac{a}{\bar{x}} + b,$$

where $\bar{x}$ is the mean trend-cycle value for a given month.

If the additive seasonal index is already computed, b is assumed to be one, and $\sum y / \sum x$ therefore can be derived by dividing the additive index for each month by the mean trend-cycle value for that month, and then adding one. If the multiplicative index is computed, the additive index can be derived by subtracting one from the multiplicative index for each month and multiplying the result by the corresponding mean trend-cycle value. If the seasonal correction factors have already been obtained from the general expression, $y_r = a + bx$, and the multiplicative index is desired, it can be derived by substituting the values of a, b, and $\bar{x}$ in the equation, and solving for $\sum y / \sum x$.

[5] The twelve-month moving average is defective as a measure of the trend-cycle component in that it reflects changes in seasonality. It does not reach sufficiently up into cyclical peaks nor down into troughs. It does not always iron out random movements of a few months' duration. It does not cover

sum of the original figures for each month is divided by the sum of the moving-average values for that month. (3) In order that the relatives given by the preceding step sum to 1,200 (or average 100), each of the twelve relatives is expressed as a percentage ratio to the arithmetic mean of the twelve relatives.

In this procedure, it is not necessary to compute the moving average.[6] If a twelve-month moving total is calculated and centered at the seventh month,[7] as is common practice, the twelve-month moving totals for the Julies are identical with the annual totals. Let a series run from y_1 to y_{49}, where y_1 is the value for January of the first year, y_2 the value for February of the first year, and so on. The moving-average values for the successive Julies are

$$\frac{\sum_{1}^{12} y}{12}, \quad \frac{\sum_{13}^{24} y}{12}, \quad \frac{\sum_{25}^{36} y}{12}, \quad \text{and} \quad \frac{\sum_{37}^{48} y}{12},$$

and the sum of the July moving-average values is

$$\frac{\sum_{1}^{48} y}{12}.$$

The moving-average values for the successive Augusts are

$$\frac{\sum_{2}^{13} y}{12}, \quad \frac{\sum_{14}^{25} y}{12}, \quad \frac{\sum_{26}^{37} y}{12}, \quad \text{and} \quad \frac{\sum_{38}^{49} y}{12},$$

and the sum of the August moving-average values is

$$\frac{\sum_{2}^{49} y}{12}.$$

But the sum of the moving-total values for the Augusts can be derived from that for the Julies by subtracting y_1 and adding y_{49}, and the sums for other months can be obtained by similar calculations.[8]

---

the full series, and it may turn as much as six months before or after the original series. These defects usually have little influence on seasonal indexes.

[6] I am indebted to Mr. Bert Gottfried for pointing out this simplification.

[7] The effect upon the seasonal index of centering the twelve-month moving average on the seventh month instead of between the sixth and seventh months is negligible.

[8] This short cut can, of course, be used to find the mean moving-average values in other problems; for example in computing the difference-from-moving-average index by the formula $\frac{\Sigma y - \Sigma x}{n}$, or in transforming an additive index to a multiplicative index.

Relation to the ratio-to-moving-average index. There is an enlightening relation between $\sum y/\sum x$ and the ordinary unweighted mean ratio-to-moving-average index, $\sum (y/x)/n$. The former is a weighted mean ratio-to-moving-average index, where the weights are the moving-average values. This is plain from the equation below, in which a single prime denotes the first year, a double prime the second year, and so on.

$$\frac{\frac{y'}{x'}\cdot x' + \frac{y''}{x''}\cdot x'' + \frac{y'''}{x'''}\cdot x''' + \cdots}{\sum x} = \frac{\sum y}{\sum x}.$$

Since the new index differs from the ordinary ratio-to-moving-average index only in that it is weighted by the moving-average values, it seems appropriate to call it the weighted ratio-to-moving-average index, or more briefly, the weighted-ratio index.

As is well known the relation between a weighted and an unweighted mean depends upon the correlation between the variates and the weights. If the Pearsonian coefficient of correlation is positive, the weighted mean will be greater than the unweighted mean; if it is negative, the weighted mean will be smaller; and if the correlation is zero, the two means will be equal. Hence, when there is little correlation between y/x and x, $\sum y/\sum x$ will tend to be a good approximation of $\sum (y/x)/n$.

A further difference between the indexes arises from the way they are centered, since the weighted-ratio index is centered by dividing each preliminary index by the average of the preliminary weighted-ratio indexes, and the ratio index is centered by dividing each preliminary index by the average of the preliminary ratio indexes.

For a given month differences in ratios are assumed to arise from random causes. In computing a seasonal index, a given random factor should receive the same weight wherever it falls. Under the weighted-ratio method, a random factor which falls at a high level receives a greater weight than a similar random factor which falls at a low level. Since the ratio method gives the same weight to all ratios, it seems superior to the weighted-ratio method in this respect.

Another advantage of the ratio method is that the ratios computed for each month can be put to a useful purpose. A chart of the ratios for successive Januaries, Februaries, and so on, is helpful in determining whether a changing or a constant seasonal pattern prevails, and what periods are appropriate in the latter case. The weighted-ratio method assumes a constant pattern and relies upon a study of the original series for the selection of periods for which this assumption seems reasonable.

A further advantage of the ratio method is that extreme items can easily be excluded from the indexes. They can be picked from the ratio charts for individual months and omitted, or one of several positional means can be selected. The elimination of extreme items is more difficult if the weighted-ratio method is used, for extreme items do not show up so clearly in charts of the original data as in charts of ratios to moving averages, and positional means based on arrays of original figures are highly suspect, particularly in series that have pronounced trends or cycles.

Relation to the monthly means index. The new index bears a close resemblance to the familiar monthly means index. This index is constructed as follows. (1) Totals of the original figures for each month are computed. (2) These totals are expressed as relatives of the arithmetic mean of the twelve totals. (3) The relatives are adjusted for trend. The trend correction can be made by computing the slope of a least-squares straight line fitted to annual averages of the monthly data expressed as relatives of their arithmetic mean and applying this slope to the relatives. The calculations are simplified by taking the origin at the middle year. The slope can then be calculated directly from the formula $b=\sum xy/\sum x^2$, where x refers to the year and y to the annual averages in relative form. Since b gives the annual increment, it must be divided by 12 to yield the monthly increment. If July is assumed to be the central point, no adjustment of the relative for this month need be made. The other months can then be corrected for trend by subtracting b' (where $b'=b/12$) from August, $2b'$ from September, $3b'$ from October, $4b'$ from November, and $5b'$ from December; and by adding b' to June, $2b'$ to May, $3b'$ to April, $4b'$ to March, $5b'$ to February, and $6b'$ to January.

The relations between the monthly means index and the weighted-ratio index may be set out as follows. Let $A, B, C, \cdots L$ refer to the months of the year, and the subscripts $1, 2, 3, \cdots n$ to the years, so that A_1 is the value for January of the first year, G_3 the value for July of the third year, L_5 the value for December of the fifth year, and so on. Assume that the moving average is available for the full period.

The uncentered weighted-ratio index for July is

$$\frac{12\sum G}{A_1+B_1+C_1+\cdots+L_n}$$

Since the monthly means method, as described here, does not require a trend correction for July, this is also the formula for the monthly means index for July.

The uncentered weighted-ratio index for August is

$$\frac{12\sum H}{B_1 + C_1 + D_1 + \cdots + A_{n+1}}$$

The monthly means index for August is

$$\frac{12\sum H}{A_1 + B_1 + C_1 + \cdots + L_n},$$

with a trend correction applied.

The difference between the two August indexes is that the monthly means index includes A_1 and not A_{n+1}, and is explicitly corrected for trend. Similar differences characterize other months. An additional, though minor, difference arises when the weighted-ratio index is centered.

The monthly means index for July may be written:

$$12\left[\frac{G_1}{A_1 + B_1 + \cdots + L_1}(A_1 + B_1 + \cdots + L_1) + \frac{G_2}{A_2 + B_2 + \cdots + L_2}(A_2 + B_2 + \cdots + L_2) + \cdots\right] \div (A_1 + B_1 + \cdots + L_n).$$

That is, the monthly means index may also be written as a weighted average of ratios. For this reason, it may be said to have the same kind of implicit weighting as the weighted-ratio index.

The weighted-ratio method assumes that the random component and most of the cyclical component in the series cancel out when the original figures for any month are summed, and the trend and the residual cycle are eliminated when the sum of the original figures is divided by the sum of the moving-average values. The latter assumption is valid when the curve used to measure the trend-cycle component is a perfect measure of it. If the twelve-month-moving-average is used in the manner suggested here, the trend-cyrle adjustment is crude, for the difference between the trend adjustments for two successive months is nothing more than the difference between the figure from the end of the series which is added and the figure from the beginning of the series which is dropped, e.g., $A_{n+1} - A_1$. This difference may be dominated by random factors. The monthly means method utilizes the slope of a straight line fitted to annual totals to adjust for the trend and residual cycle. This method has the advantage that the trend adjustment is

more stable from month to month, but it suffers from all the difficulties connected with fitting a trend, and in addition from the fact that it adjusts only for a linear trend. On theoretical grounds alone it is difficult to say that one of these trend adjustments is better than the other.[9]

A disadvantage of the new method is that it involves the loss of a half year at each end of the series. The monthly means method must be confined to complete years, but this does not involve so great a loss, since seasonal indexes can be computed on a fiscal-year basis. This disadvantage of the weighted-ratio index is not very serious, partly because most series are divided into several seasonal periods, and only the first and last periods are affected, and partly because the moving average can be made to cover the full period by free-hand extrapolation.

Comparisons of indexes for particular series. Seasonal indexes have been computed by the weighted-ratio method, the monthly means method, and the ratio method for ethyl alcohol production in the United States, 1921–40, refined copper shipments in the United States, 1927–36, and Douglas fir production in the United States, 1918–27, and time records have been kept. Table I shows that the weighted-ratio method

TABLE I

TIME REQUIRED TO COMPUTE MULTIPLICATIVE SEASONAL INDEXES BY DIFFERENT METHODS

Method	Ethyl alcohol production, 1921–40		Refined copper shipments, 1927–36		Douglas fir production, 1918–27	
	Minutes	Per cent of time for weighted-ratio method	Minutes	Per cent of time for weighted-ratio method	Minutes	Per cent of time for weighted-ratio method
Monthly means	57	119	42	108	34	110
Ratio	161	335	97	249	95	306
Weighted ratio	48	100	39	100	31	100

took the least time in each case. In this respect its advantage over the ratio method is considerable, and its advantage over the monthly means method small but persistent. It must be pointed out, however, that the time advantage over the monthly means method is not conclusive, since the monthly means method will be somewhat shorter if the linear trend

[9] It is worth pointing out that the difference-from-twelve-month-moving-average index is not, as is nearly universally assumed, obviously superior to the additive monthly means index. The difference-from-moving-average index for a given month is $\frac{\Sigma(y-x)}{n}$, or $\frac{\Sigma y - \Sigma x}{n}$, and the term which adjusts for trend is Σx. Since this is precisely the term which adjusts for trend in the weighted-ratio index, the difference-from-moving-average index, like the weighted-ratio index (and unlike the ratio index), involves a crude trend correction.

is derived by the method of selected points instead of least squares; though this procedure may yield a cruder index.

Seasonal indexes have also been computed by each of the three methods for seven additional series.[10] The centered indexes for the ten series are given in Table III. Table II compares the indexes by the device of two summary measures. The first, the coefficient of similarity, is computed as follows. (1) The deviations from 100 of each of the pair of seasonal indexes under comparison are computed. (2) One set of deviations is adjusted so that its sum is equal to the sum of the second set, which is considered the base. (3) The differences between the adjusted deviations, month for month, are summed without regard to sign. (4) Finally the sum of the differences is divided by the base figure and one subtracted from this quotient.[11] Since this measure eliminates differences in amplitude, the table gives another measure which

TABLE II

COMPARISON OF SEASONAL INDEXES COMPUTED BY DIFFERENT METHODS

Series	Coefficient of similarity			Average difference		
	Weighted-ratio and ratio index	Monthly means and ratio index	Monthly means and weighted-ratio index	Weighted-ratio and ratio index	Monthly means and ratio index	Monthly means and weighted-ratio index
1. Steel ingot production in U.S., 1921–38	.992	.895	.956	0.54	0.58	0.35
2. Refined sugar shipments in U.S., 1927–39	.887	.849	.899	0.42	1.01	1.77
3. Douglas fir production in U.S., 1916–27	.933	.912	.941	0.53	0.68	0.35
4. Ethyl alcohol production in U.S., 1931–40	.990	.970	.972	0.78	0.37	0.72
5. Crude cottonseed oil production in U.S., 1916–24	.977	.985	.970	1.53	1.06	1.82
6. Cotton consumption in U.S., 1914–22	.926	.897	.858	0.18	0.60	0.47
7. Cotton consumption in U.S., 1923–39	.966	.944	.948	0.29	0.51	0.28
8. Corporate issues of stocks in U.S., 1907–27	.690	.655	.921	4.22	5.05	1.25
9. New corporate issues of bonds and stocks in France, 1922–38	.584	.520	.955	12.53	12.50	1.20
10. London bank clearings, 1898–1913	.934	.902	.918	0.24	0.33	0.22

[10] Strictly comparable procedures have been used in computing all ten indexes. Periods for which the moving average is fully available were selected and no attempt was made to omit extreme items.

[11] See Simon Kuznets, *Seasonal Variations in Industry and Trade* (National Bureau of Economic Research, 1933) pp. 281–282.

TABLE III

SEASONAL INDEXES COMPUTED BY THE MONTHLY MEANS METHOD, RATIO-TO-MOVING-AVERAGE METHOD, AND THE WEIGHTED RATIO-TO-MOVING-AVERAGE METHOD

Series	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Total
1. Steel ingot production in U.S., 1921-38													
Monthly means method	99.2	106.0	102.7	105.0	105.6	95.2	94.3	97.6	97.1	99.1	94.6	90.8	1200.2
Ratio method	99.3	106.6	102.1	103.2	106.6	99.2	94.2	96.6	96.4	99.4	94.8	89.7	1200.1
Weighted-ratio method	99.1	106.3	110.0	105.4	106.1	98.4	94.5	97.6	96.9	98.7	93.9	89.9	1199.8
2. Refined copper shipments in U.S., 1927-38													
Monthly means method	95.9	98.2	101.9	97.3	98.2	102.4	95.1	102.4	102.1	111.9	101.2	93.5	1200.1
Ratio method	93.0	97.8	103.4	97.7	98.9	103.8	98.1	101.8	100.8	110.5	98.4	89.8	1199.8
Weighted-ratio method	95.7	97.7	104.0	98.9	99.2	102.8	98.0	101.7	100.7	109.7	98.3	90.0	1199.7
3. Douglas fir production in U.S., 1919-27													
Monthly means method	87.2	95.9	101.9	101.7	105.9	107.8	91.5	105.0	104.7	105.0	102.6	84.9	1199.9
Ratio method	86.0	97.6	100.7	101.5	106.3	108.1	91.8	105.8	103.6	105.4	103.1	85.4	1200.3
Weighted-ratio method	86.7	95.7	101.6	101.2	105.6	107.4	91.5	105.3	103.0	105.3	103.0	85.8	1200.1
4. Ethyl alcohol production in U.S., 1921-40													
Monthly means method	93.8	83.1	87.1	84.6	90.0	91.0	97.8	104.0	109.0	126.4	118.5	114.1	1200.1
Ratio method	94.9	82.9	86.4	84.7	90.2	91.0	97.1	104.4	109.4	126.7	118.0	113.8	1200.1
Weighted-ratio method	94.9	83.7	87.7	85.2	90.6	91.5	97.9	104.8	108.6	125.5	117.3	112.4	1200.1
5. Crude cottonseed oil production in U.S., 1918-24													
Monthly means method	161.0	129.1	103.8	62.4	39.1	21.3	11.9	15.3	80.8	199.4	206.1	169.8	1200.0
Ratio method	163.6	131.0	104.0	61.4	38.0	20.4	11.4	15.0	80.9	200.8	206.0	167.4	1200.1
Weighted-ratio method	164.4	132.0	106.1	63.7	39.8	21.5	11.6	14.5	79.0	196.6	203.5	167.2	1199.9
6. Cotton consumption in U.S., 1914-22													
Monthly means method	103.7	95.7	105.5	100.2	103.7	102.1	97.6	99.1	96.8	99.7	99.3	96.7	1200.1
Ratio method	104.2	96.6	106.0	100.8	104.0	102.5	97.8	98.8	96.6	99.0	98.1	95.4	1199.8
Weighted-ratio method	104.4	96.4	106.0	100.7	104.0	102.2	97.6	98.9	96.5	99.1	95.5	95.8	1200.1
7. Cotton consumption in U.S., 1923-39													
Monthly means method	106.6	99.7	109.1	102.1	101.7	95.5	88.9	95.1	96.7	105.9	102.7	96.0	1200.0
Ratio method	107.5	100.6	109.8	102.4	101.6	94.9	88.4	94.4	96.4	106.2	102.4	95.5	1200.1
Weighted-ratio method	107.1	100.2	109.5	102.3	101.7	95.5	88.8	94.8	96.4	105.6	102.3	95.6	1199.8
8. Corporate issues of stocks in U.S., 1907-27													
Monthly means method	110.7	113.0	107.4	99.7	84.4	108.0	82.0	85.1	73.7	95.3	92.4	148.3	1200.0
Ratio method	111.0	116.6	104.3	105.8	86.0	115.5	87.6	90.6	66.6	86.2	92.5	137.3	1200.0
Weighted-ratio method	111.0	113.7	108.4	100.7	85.0	109.9	83.5	86.0	74.2	94.8	91.4	142.3	1200.0
9. New corporate issues of bonds and stocks in France, 1922-38													
Monthly means method	110.0	168.7	82.5	93.3	99.5	139.5	70.3	39.4	61.6	148.3	92.2	94.9	1200.2
Ratio method	108.6	132.7	77.6	110.2	119.3	118.6	78.0	44.5	67.3	139.2	112.2	91.7	1199.9
Weighted-ratio method	109.4	165.3	81.6	91.7	99.6	139.1	70.2	38.8	62.7	152.3	93.5	95.9	1200.1
10. London bank clearings, 1888-1913													
Monthly means method	104.3	102.9	101.1	105.9	100.4	98.7	102.6	94.5	90.3	99.4	97.1	102.9	1200.1
Ratio method	104.5	103.3	101.8	105.9	100.4	98.8	102.7	95.0	90.3	98.8	96.7	102.0	1200.2
Weighted-ratio method	104.7	103.1	101.4	106.2	100.5	98.8	102.6	94.5	90.1	99.1	96.9	102.4	1200.3

takes amplitude into account—the average difference disregarding sign. The average difference is computed by subtracting one index from another, month for month, summing without regard to sign, and dividing by 12.

Tables II and III support the conclusion that there is a greater similarity between the monthly means index and the weighted-ratio index than there is between either of these indexes and the ratio index. The coefficients of similarity between the monthly means and the weighted-ratio indexes are highest in six of the ten series and the average difference is smallest in six of the ten series. The greater resemblance is most notable in the two series which contain a large random component: corporate issues of stocks in the United States, 1907–27, and new corporate issues of bonds and stocks in France, 1922–38. In such series the implicit weighting exerts its greatest effect and substantial differences between the weighted-ratio indexes and the ratio index may be expected.

The tables also suggest that the weighted-ratio index provides a better estimate of the ratio index than does the monthly means index. The new index is closer to the ratio index in eight series according to the average difference and in six series according to the coefficient of similarity. In no case is the monthly means index appreciably closer to the ratio index than is the weighted-ratio index, whereas in refined copper shipments the weighted-ratio index is much closer to the ratio index.

Summary. The weighted-ratio index has some advantages over the monthly means index. It can be computed more rapidly, it is a special case of a general method, and it can be converted to a strictly analogous additive index. The empirical evidence suggests that the weighted-ratio index gives a better approximation to the ratio index. But in the absence of strong theoretical support, the margin of superiority is too small to warrant a generalization.

Like the weighted-ratio index, the monthly means index is a centered average of ratios weighted by trend-cycle values. The two indexes almost always come out very nearly the same. They usually give close approximations to the ratio index. But if the random component or the correlation between the ratios and the corresponding trend-cycle values is large, the weighted-ratio indexes may depart considerably from the ratio index.

SEASONAL FACTORS DETERMINED BY DIFFERENCE FROM AVERAGE OF ADJACENT MONTHS

By Henry A. Latané
Lionel D. Edie and Company, Inc.

This method for smoothing economic time series is based on a comparison of the data for each individual month with the average of such data for the two adjacent months. It is assumed that, in any time series from which variations due to seasonal factors have been satisfactorily eliminated, the index for any one calendar month is typically equal to the average of the indices of the two adjacent months. For example, the adjusted February indices, over a series of years and after eliminating data obviously affected by extraneous factors such as strikes, should fluctuate around the averages of the adjacent January and March indices and typically should equal such averages. Adjustment factors which produce such results are taken to be satisfactory seasonal factors.

The method is based on the use of moving averages and so is directly comparable with other such methods now in use. Its chief advantage is that the comparison is close to the data and consequently there are fewer extraneous factors to hide the seasonal pattern. This is illustrated by showing the per cent by which railroad freight revenues[1] for each month over a series of years differs (1) from the average of the adjacent months and (2) from a twelve-month centered moving average. Examples are:

	April		July	
	(1) Per cent	(2) Per cent	(1) Per cent	(2) Per cent
1936	.4*	−3.4	3.6	1.8
1937	−2.4	−1.7	3.4	3.0
1938	−3.4	−11.0	3.6	.4
1939	−6.8*	−13.7*	3.2	−1.8
1940	−2.4	−9.6	2.6	1.0

* Comparison affected by coal strike.

When the adjacent months are used as a base for comparison the range of variation within each calendar month is consistently smaller and the effects of random events are more clearly indicated. There are seven calendar months wherein the range among the percentage differences

[1] Class I Railway Freight Revenues (I. C. C. data) adjusted to a uniform month basis but not adjusted for seasonal variation.

(from the average of adjacent months) is less than 2 percentage points, after eliminating data affected by known extraneous factors. The widest such range is nearly 7 points, but, even so, it is smaller than the range for all but two months based on the twelve-months-moving average.

The greater consistency simplifies the problem of selecting the proper amount of variation to be eliminated. The shorter base also permits use of current data earlier than is possible with other averaging techniques. Using a centered average avoids the problem of eliminating trend and thus escapes one of the major difficulties in the link-relative method. The time and skill involved in applying the average of adjacent months is approximately the same as that of other simple averaging techniques.

A corollary but important advantage of the method is that it furnishes a direct check on the work-day adjustment factor. This is of especial significance in time of war when the work week for many industries is unknown or changing. Take the work-day adjustment for railroad freight revenues as an example. The Federal Reserve Board uses a six-day week for adjusting these data. In their adjusted series for the period from July 1935 through July 1941, twenty months with five Sundays were above the average of the adjacent months, while only four were below such average. Of the "four-Sunday" months, only eleven out of forty-one were above the corresponding average. It therefore is obvious that too much allowance was made for the "five-Sunday" months. The median index for such months was 2.1 per cent above the average of the two adjacent months, indicating that adjustment for a week of 6½ working days would be considerably more satisfactory than the six-day week actually used by the Board. The same series adjusted on a 6½ day basis showed only 13 out of 24 "five-Sunday" months above the adjacent months average. While there is little actual origination of freight traffic on Sunday, enough is received from connections to make a 6½ day working week entirely logical.

The average of adjacent months is valuable for checking whether or not seasonal variation has been successfully eliminated in any given series. When a test shows that the adjusted indices for any one calendar month are consistently above or below the average of adjacent months, it is a strong indication that further adjustment for seasonal is needed. If constant seasonal factors were used in the original computations, it is possible to re-compute these factors by the adjacent months method directly from the adjusted data. The adjusted Federal Reserve Board series for railroad freight revenues was reworked in this manner and new seasonal factors computed which differed very materially from those used by the Board but which showed a maximum difference of

.3 per cent from factors obtained by working directly from the original data.

The proposed method also escapes the heavy dependence on judgment which characterizes non-mathematical techniques. It is particularly valuable for determing the net difference of two series such as revenues and expenses where it is very important to keep the method for eliminating seasonal variations uniform. It was for studies of railroad net income that this method was originally developed. Here the difference between revenues and expenses is small and seasonal adjustments must be refined to be of any value. The method has been in use for this purpose for several years.

Table 1 shows railroad freight revenues adjusted by this method as compared with the corresponding Federal Reserve Board series. As is apparent, the differences are substantial. In another more recently adjusted Board series, used as a test,[2] seasonal factors derived by the average of adjacent months method differed by no more than .5 per cent from those of the Board.

In deriving seasonal factors by the proposed method it is necessary to determine (1) the per cent by which the data for each calendar month typically differs from the average of the adjacent months and (2) the seasonal factors necessary to eliminate this typical difference. The typical difference may be determined by averaging the months under consideration after eliminating data known to be affected by extraneous factors. The percentage variation from adjacent months[3] for each April, for example, from 1936 through 1940 has already been given. It is known that data for April 1936 and 1939 were affected by coal strikes. The average of the other three months is taken as the typical difference. None of the corresponding figures for July were affected by known extraneous factors so an average of all of these items was used. Determination of the months to be eliminated is partly a matter of judgment,

[2] By Mr. R. C. Barton, Jr., the Division of Research and Statistics, the Federal Reserve Board. Mr. Barton's comment was "In the one experiment which we were able to make using your method on a series which had been recently adjusted (in contrast to freight revenues which we know is bad) we found that it yielded a somewhat smoother adjustment but in no case were the factors for a particular month more than .5 per cent apart." Mr. Barton was consulted as an authority on seasonal factors in economic data and his generous encouragement led to preparation of this material for publication.

[3] The arithmetical derivation of these percentage variations is simple. For example,

	Original Data	Take Difference	Take Difference	Halve	Divide by Original Data
January	90				
February	100	10	20	10	10%
March	90	−10	0	0	0
April	80	−10			

TABLE I

MONTHLY FREIGHT REVENUES, CLASS I RAILROADS

(millions of dollars)

	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
	(1) Data adjusted for working days but not for seasonal variation											
1935	209.7	221.9	228.3	226.4	219.0	228.9	219.1	235.3	252.2	273.0	235.7	230.7
1936	240.0	257.3	249.0	251.8	261.2	266.9	257.2	277.0	[illegible]	312.1	313.1	297.1
1937	271.3	297.5	300.0	285.0	265.3	270.2	261.5	281.9	302.0	299.1	265.9	273.3
1938	220.3	216.0	217.3	209.4	215.2	220.7	200.1	242.7	268.0	285.0	271.8	219.1
1939	249.9	244.2	216.1	258.2	237.0	253.2	261.2	284.3	323.0	315.8	315.9	271.1
1940	281.1	270.3	260.5	252.4	276.0	284.0	276.9	298.1	302.0	331.2	325.0	312.2
1941	299.9	322.2	337.5	302.5	351.8	390.2	391.0	399.5	427.6	421.2	403.3	385.7
1942	389.1	410.8	432.9	463.8	463.6	459.6	426.3					
Proposed Seasonal Factors	95.1	94.8	96.1	92.5	95.4	98.3	102.1	109.9	110.5	111.0	105.6	95.5
	(2) Data adjusted for seasonal variation by the proposed method											
1935	220	237	232	245	230	232	223	223	231	252	211	245
1936	252	271	259	272	274	277	250	254	272	285	296	311
1937	285	303	312	308	278	290	245	257	275	277	251	240
1938	232	228	226	221	227	223	225	249	242	251	257	261
1939	261	258	258	312	212	253	243	252	292	303	301	[illegible]
1940	298	285	271	282	292	295	230	298	300	298	307	327
1941	323	340	351	323	379	396	391	392	387	379	381	401
1942	409	433	452	496	507	510	418					
F R B Seasonal Factors	95	97	99	97	98	97	101	99	106	111	103	95
	(3) Data adjusted for seasonal variation by Federal Reserve Board											
1935	219	232	228	229	221	236	215	225	242	244	217	252
1936	250	265	251	261	270	272	272	283	277	279	204	311
1937	289	296	300	291	271	286	265	257	276	272	257	210
1938	235	223	217	211	225	220	211	213	217	260	242	240
1939	265	251	215	236	239	259	266	284	296	315	309	291
1940	291	278	264	270	281	297	291	298	313	298	311	333
1941	321	333	315	310	365	396						

but not usually difficult. Variation in coal production is the most important single factor influencing railroad freight revenues in normal times, and the effects of the numerous bituminous and anthracite coal strikes in recent years are obvious.

The second step is to determine the seasonal factors which will eliminate the typical variation of each month from the average of adjacent months. The mathematical solution of this problem involves a set of twelve simultaneous equations. Fortunately, however, these equations can be reduced to the first degree and solved by simple cumulative addition, as follows:

Let:

$a_1, a_2 \cdots a_{12}$ be the typical percentage differences between each

month and the average of the adjacent months. The numerical values for the a's in the illustrative example are given in Table II,

TABLE II

MONTHLY FREIGHT REVENUES

	a_1 etc.	(b)[a] $2M_1a_1$ etc.	(e)[a]	(g)[a]		Per cent difference M_1 etc. from R	Seasonal factors
						4.0	95.1
1	.10			.4	.4	5.2	94.8
2	− .75	− 1.7	−1.7	−1.3	− .9	3.9	95.1
3	2.05	3.9	2.2	2.6	1.7	6.6	93.5
4	−2.30	− 4.6	−2.4	−2.0	− .3	4.0	95.4
5	.75	1.2	−1.2	− .8	− 1.1	3.7	96.3
6	−2.40	− 4.9	−6.1	−5.8	− 6.9	− 2.1	102.1
7	3.45	6.8	.7	1.1	− 5.8	− .9	100.9
8	−5.13	−10.6	−9.9	−9.5	−15.3	−10.5	110.5
9	2.90	6.0	−3.9	−3.5	−18.8	−14.0	114.0
10	5.80	11.7	7.8	8.2	−10.6	− 5.8	105.8
11	1.15	2.1	9.9	10.3	− .3	4.5	95.5
12	−5.10	− 9.9	− .	.3	—	4.0	
			−4.6 (f)[a]		−57.9 (i)[a]		
			− .3 10/12		− 4.8 3/12		

[a] Letters refer to equations used in text.

$M_1, M_2 \cdots M_{12}$ be arbitrary values for freight revenues for January, February, $\cdots$ December, satisfying the conditions of the a's,

R equal average monthly freight revenues for the whole period.

Then:

The problem is to determine the M's numerically in relation to R as this will give the desired seasonal factors.

The M's are determined numerically from the twelve simultaneous equations of the form:

$$\text{(a)} \qquad \frac{M_1 - \dfrac{M_{12} + M_2}{2}}{M_1} = a_1 \text{ etc.}$$

or, restating set (a)

$$\text{(b)} \qquad \begin{aligned} (M_1 - M_{12}) - (M_2 - M_1) &= 2M_1a_1 \\ (M_2 - M_1) - (M_3 - M_2) &= 2M_1a_2 \text{ etc.} \end{aligned}$$

To solve these equations, arbitrary values must be assigned to the M's used as factors on the right side. The values for the M's differ by only a relatively small percentage from the average freight revenues for the period; therefore the error is not large if R is substituted for

these M's. Solving the equations after this substitution gives an approximation to the various M's in terms of R. The equations are then reworked substituting the M's so determined on the right side. This gives values for the M's which can be used to determine the seasonal factors.

It must be emphasized that this method of solving the equations introduces only a negligible error in the final figure. Seasonal factors derived from the first approximation (using R for all M's on the right side) show a maximum difference of 1.6 points from those derived from the second approximation but factors obtained from a third approximation showed a maximum deviation of only .2 percentage points. The second approximation is accurate enough for all practical purposes and is used in this study.

Substituting numerical values derived as explained above in the right side of set (b) the equations become:

(c) $(M_1-M_{12})-(M_2-M_1)=.00\ R$ etc.
Letting $d_1=M_1-M_{12};\quad d_2=M_2-M_1;\cdots\quad d_{12}=M_{12}-M_{11}$
set (c) becomes:

(d) $d_1-d_2=.00\ R;\ d_2-d_3=-.017\ R;\cdots d_{12}-d_1=-.099\ R.$
By cumulative addition, this set of equations becomes

(e) $d_1-d_2=.00\ R;\ d_1-d_3=-.017\ R;\cdots d_1-d_{12}=+.099\ R$
$d_1-d_1=0.$
Adding set (e):

(f) $12d_1-(d_1+d_2\cdots d_{12})=-.046\ R$
but $d_1+d_2\cdots d_{12}=(M_1-M_{12})+(M_2-M_1)\cdots(M_{12}-M_{11})=0$
so: $d_1=-.003\ 10/12\ R.$
Substituting for d_1 in set (e):

(g) $-d_2=.004\ R;\ -d_3=-.013\ R$ etc.
but $-d_2=M_1-M_2$ etc.

(h) so $M_1-M_2=.004\ R$ etc.
which is in form corresponding to set (d) and can be reduced to:

(i) $12M_1-(M_1+M_2\cdots M_{12})=-.579\ R$
but $(M_1+M_2\cdots M_{12})=12\ R$

(j) so $M_1=R-.048\ 3/12\ R.$
Numerical values of the other M's may be readily obtained.

Table II shows the full numerical derivation of each set of equations. Seasonal factors which minimize the differences of the various M's from R were taken as final and were used to smooth the series.

MOVING SEASONAL INDEXES

By Dudley J. Cowden
University of North Carolina

A widely used method of computing a seasonal index, whether a stable index or a moving seasonal, is based upon percentages of a twelve month moving average. The moving average is intended as a rough estimate of combined trend and cyclical movements. The moving average, however, tends to smooth out cyclical peaks and troughs. As a consequence percentages of moving average have a downward bias in years containing a sharp trough of depression and an upward bias in years containing a sharp peak of prosperity. In computing a stable seasonal the inaccuracy occasioned by this lack of homogeneity in the data is partly overcome by using a modified mean or median for the monthly arrays.

A moving seasonal is usually obtained by fitting a free hand trend to the percentages for each month separately, and adjusting the trend values so that the total for each year will be 1200, at the same time retaining the smoothness and good fit of each trend. The method is both laborious and subjective, and suffers even more than does a stable seasonal from the lack of homogeneity of the data.

If the percentages of moving average are adjusted (by multiplying by a constant) so that they will total 1200 for each year two objects will be accomplished: (1) the bias due to inflexibility of the moving average will be largely overcome; (2) if polynomial trends are fitted by the method of least squares to each month, the trend values will total 1200 for each year, thus avoiding the adjustment process referred to above. This second result will hold only if the same degree of equation is used for each month.

The reason the trend values total 1200 is that, if an equation of the type $Y_c = a + bX + cX^2 \cdots$ is used, $\Sigma a = 1200$, $\Sigma b = 0$, $\Sigma c = 0$, etc. If one merely wishes to have his seasonal index total 1200 each year, he need not adjust the percentages of moving average to total 1200 before constructing his index. All he need do is to adjust the 12 values of a so that they will total 1200, and the twelve values of each other constant so that they will total zero. This, however, does not overcome the difficulty occasioned by the defect inherent in the moving average, and sometimes the adjustment of the constants will even change the sign of certain constants.

If a time series is deseasonalized by dividing by a seasonal index constructed as described above, the total of the deseasonalized data will

not, in general, be the same as the total of the original data. It would be possible to overcome this difficulty by constructing the seasonal index from differences between the original data and the 12 month moving average, and to deseasonalize by subtracting the seasonal index from the original data. This procedure would be logical if one thinks of the original data as being the sum of the component elements: Trend (T); Cycle (C); Seasonal (S); Irregular (I). According to this concept seasonal movements vary in absolute magnitude according to the trend values and the position of the cycle. A stable seasonal is virtually precluded, however, and moving seasonals are described by unnecessarily complex trends. Consistent with the above concept, but more easy to handle statistically, is to think of the original data as consisting of $T \times C' \times S' \times I'$, the primes indicating that the movements in question are percentages of the combined preceding elements. Also, log original data $=\log T+\log C'+\log S'+\log I'$. If now a logarithmic seasonal index totaling zero is based on the differences between the logarithms of the original data and the moving average of those logarithms, and if seasonal adjustment is made by subtracting the logarithms of the seasonal index from the logarithms of the original data, then the total of the deseasonalized logarithms will be the same as the total of the original logarithms, and the geometric mean of the deseasonalized data will be the same as the geometric mean of the original data. (Also, the geometric mean of the seasonal index will be unity, and the geometric mean of the data deseasonalized by dividing by the seasonal index would be the same as the geometric mean of the original data.) In constructing a logarithmic seasonal index, the difference between the logarithms of the original data and the moving average of those logarithms should be adjusted to total zero each year before computing the seasonal index. If this is done, the total of each of the different trend constants will also total zero.

For the method of computing a moving seasonal index described here, use of orthogonal polynomials will greatly facilitate the computation of the constants, and also the use (with a grain of salt) of covariance analysis for testing the degree of equation to use.

ERRORS IN CARD PUNCHING

By W. Edwards Deming, Benjamin J. Tepping, and Leon Geoffrey
Bureau of the Census

The purpose of this study is to obtain information on errors in card punching, the ultimate object being (i) to provide a basis for determining how such errors, if uncorrected, would affect the Census tabulations, and (ii) to provide a basis for administering sample verification. The introduction of the sample verification[1] of coding and card punching in the Census, and its adoption and extension in large corporations, renders such studies needful. Naturally, this paper must deal with the processing of data in the Census, but many of our conclusions are applicable in other lines of work.

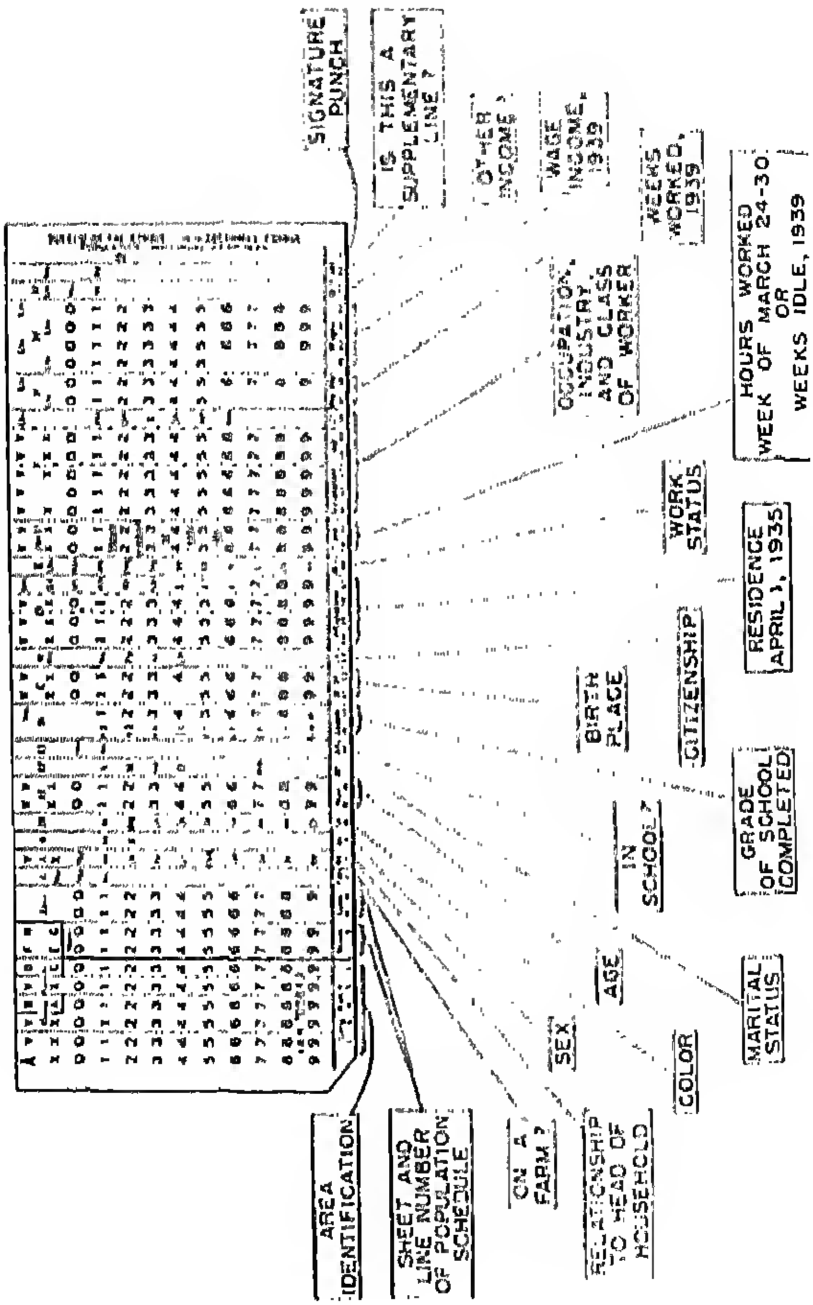
Data collected during a Census enumeration go through many stages of processing before they are summarized in the form of tables for publication. A major stage is the transfer to punch cards of the edited and coded information that appears on the original Census schedules, and on transcription sheets made therefrom. Even the most reliable punch operator occasionally punches a hole in the wrong position on a card, or skips a column, or puts a second hole in a column where there should be only one. It is with these errors that this paper deals.

After punching, the cards are sent through the verifying operation. The operator of a verifying machine is expected to discover any error in a card that goes through her machine, whereupon the card is sent back to the original operator who is obliged to punch a new one. Depending on her skill, a punch operator is thus obliged to replace from .5 per cent to 5 per cent of the cards that she punches. The verification and repunching is therefore an item of considerable expense, and any information that will help diminish it should pay high dividends.

The 25,000 study cards. The cards studied were wrongly punched A cards of the 1940 Census of Population. An A card (Chart I) was punched for each person enumerated in the 1940 Census. Five samples, each consisting of 5,000 wrongly punched A cards, were studied. Tests showed that the study cards were sufficiently representative of all wrongly punched A cards to render valid the conclusions drawn in this paper.

[1] See W. Edwards Deming and Leon Geoffrey, "On Sample Inspection in the Processing of Census Returns," this Journal, Vol. 36, September 1941, pp. 351–360. The procedure for sample verification of punch cards, and the extent to which it was used to supplant 100 per cent verification, is described there. Sample verification was introduced only for certain steps in the processing of the Census data, and even then, only for clerks whose work was known individually to be uniformly so good that complete verification and correction of errors would constitute needless and wasteful refinement of their work.

CHART I
THE POPULATION A CARD



THE EFFECT OF WRONG PUNCHES WITHIN SPECIFIC FIELDS OF THE POPULATION A CARD

The effect of wrong punches on the tabulated results. The purpose of this study is to learn how sample verification of card punching affects the tabulations. In sample verification,[1] punchers having a uniform error rate below 1 wrong card per 100 cards punched were "qualified"—i.e., only 5 per cent of their work was put through the verifying machine. The wrong cards in the other 95 per cent went through to tabulation uncorrected. The fact is that even in 100 per cent verification, some errors do get through. The only way to get good cards to tabulate is to have them punched correctly in the first place. Sample verification actually raises the quality of punching by putting a premium on good punching. In the punching of the Population A card (Chart I), and the Housing E card (not shown), at any one time not more than a third of the punchers were "qualified"; and with careful controls used, the qualified punchers contributed no more than their proportionate share of the wrong cards tabulated. Moreover, the study here described shows that when sample verification is used, the net effect of wrong punches, as they occur in practice, is often negligible. There are two reasons for this. First, in many fields of the card (see Chart I), errors tend to compensate for one another. Thus, while some cards were punched male when they should have been punched female, approximately an equal number were punched conversely, and the net error in the tabulations would have been small, even had none of the wrong cards been corrected. Second, it was found that the right-hand columns of a field (the units) are oftener wrongly punched than the earlier ones (tens or hundreds), so that in fields in which a numerical quantity is punched, the errors are mostly of small magnitude. The same pattern fortunately appears in punching fields of two and three digit codes; the right-hand columns are punched wrong oftener than the left-hand ones. The reasons behind this phenomenon are undoubtedly psychological.

Wrong punches in the age field. For cards sample verified, the maximum error rate is *1 wrong card in 100 cards punched.* For the cards verified 100 per cent, the proportion of wrong cards sent to tabulation is less. But even if the average error rate is 1 wrong card in 100, wrong punches in the age field would bias the tabulated results very little even if no wrong cards were corrected. Age was incorrectly punched on 3,210 cards among the 25,000 wrongly punched cards. However, 41.8 per cent of these 3,210 were still in the correct 5-year age group, and of the remaining 58.2 per cent, 2 out of 5 were in the adjoining 5-year age groups. On this basis we may make a prediction concerning the maxi-

mum error in the final tabulations arising from sample verification. Of the 25,000 study cards, 1,868 were punched in the wrong 5-year age group. At the rate of 1 wrong card in 100, the 25,000 study cards are all the wrong cards there are for 2,500,000 people, hence only 1,868/2,500,000 or 0.075 per cent of all ages would be classified in the wrong 5-year age group.

This conclusion holds for any cross-tabulation into which age enters. If, however, one is interested only in the distribution of a population by age, the distortion introduced by these errors is even smaller. Thus 428 of the 25,000 cards should have been punched "0 to 4 years," while actually 487 were so punched, a difference of 13.8 per cent. The average amount by which the frequency in any 5-year age group differed from the correct frequency, for the 3,210 cards that were wrongly punched in the age field, was 14.3 per cent. The error introduced in a 5-year age group of the minimum of 2,500,000 cards from which the 25,000 cards are to be considered drawn will thus be less than 0.02 per cent.

Wrong punches in other fields of the card. The pattern of wrong punches in the wage income field is similar to that in the age field. Other fields studied were those for farm residence, color, citizenship, and marital status. The picture is essentially the same for all these fields. Most of the cards were correctly punched, but among the few cards that were wrongly punched, there was a tendency for the errors to compensate for one another. Thus, 24,148 of the 25,000 wrongly punched cards were correctly punched with regard to color; for the 852 cards that were wrongly punched in that field, the distribution by color was not greatly different from what it would have been without the errors.

FACTORS ASSOCIATED WITH THE OCCURRENCE OF WRONG PUNCHES

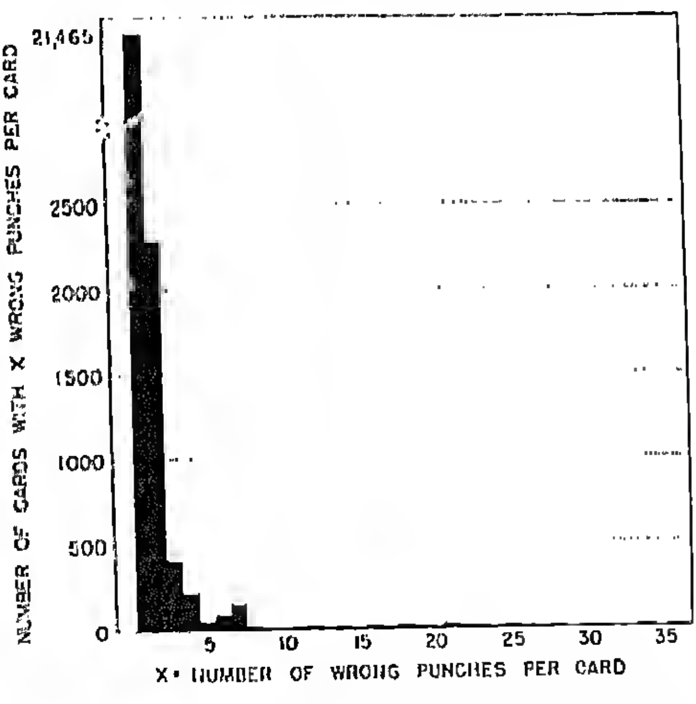
Clustering of the wrong punches on cards having more than one wrong punch. A "wrong card" has one or more wrong punches. The "error rate" of a puncher is defined as the number of wrong cards per hundred cards punched. Although this method of counting errors has proved administratively convenient, it reveals nothing about the number of wrongly punched columns. It is therefore of interest to examine cards that have wrong punches, in order to determine how many of them have only one wrong punch, how many have two wrong punches, etc.

There are 45 columns on a Population A card (Chart 1). The machine operated by the puncher is set so that the punching begins with col. 7 and ends with col. 45. The first six columns, which identify the cards by

geographic area, are later punched mechanically. On each card, therefore, there are 30 possibilities for making an error.

Chart II shows the distribution of the 25,000 study cards according to the number of wrong punches per card. Of the wrong cards 86 per cent have only one error, 9 per cent have exactly two errors, and 5 per cent have three or more errors. The average number of wrong punches per wrong card for the 25,000 cards is 1.47. The five subsamples, consisting of 5,000 cards each, exhibited stability about the same value, the averages being 1.51, 1.39, 1.47, 1.54, and 1.42.

CHART II
THE DISTRIBUTION OF THE 25,000 STUDY CARDS ACCORDING TO THE NUMBER OF WRONG PUNCHES PER CARD



Of the 25,000 wrongly punched cards, 3,535 cards contained more than one wrong punch each. An investigation of these "multiple-error" cards showed that often the wrong punches on a card were related. For example, the operator (possibly through fault of the machine) may skip col. 21, whereupon she punches the code for col. 22 into col. 21, the code for col. 23 into col. 22, etc., with the result that all the entries to the

right of col. 20 are wrongly punched, all as the result of the one wrong punch in col. 21.

Multiple errors may be classified as follows:

i. Failure to punch a code belonging in a given column of the card. As a result, each succeeding punch is entered one column to the left of its correct position. This does not mean, however, that every succeeding column is wrongly punched, since the correct punch often appears by coincidence or as a result of a compensating wrong punch.

ii. Failure of the punching machine to move, causing two punches in a given column. This has the same effect as i.

iii. The same code punched in two successive columns. As a result of this, each succeeding punch is entered one column to the right of its correct position.

iv. Interchange of punches. These are instances where the operator has, for example, punched 32 instead of 23, or 545 instead of 454.

v. Wrong punches not related to each other.

Each of the first four types of error results in a card that contains several related wrong punches.

Most errors of types i, ii, and iii are easily identified. On a large proportion of the cards containing errors of types i and ii, the punching ends with col. 44 instead of with col. 45. Such cards therefore have no "signature" punch, which the punching instructions specify must be punched in the "2" position of col. 45 of every card (see Chart I). On the other hand, most cards containing errors of type iii have two punches in col. 45, namely, the punch belonging to col. 44 and the "signature" punch.

The importance of these observations lies in the fact that cards with errors in col. 45 can be separated from the other cards at some stage of the sorting process. Such cards contain many more errors than the average, as is indicated by Table I, which shows the average number of wrong punches per card for various groupings of the 25,000 error cards.

TABLE I

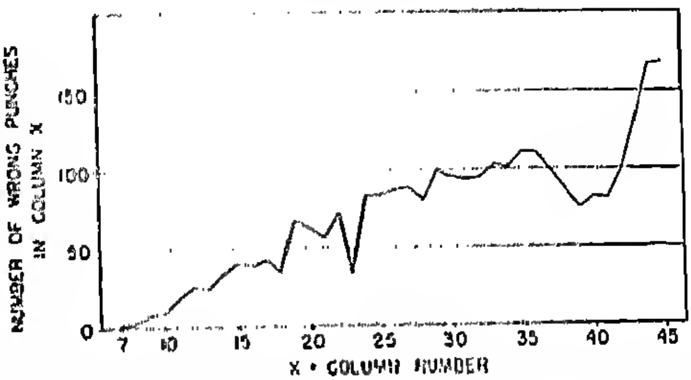
Grouping	Average number of errors per card
25,000 error cards	1.47
3,635 multiple-error cards	4.29
108 cards with no signature punch in col. 45	10.48
24,892 cards with a signature punch in col. 45	1.30

The signature punch in col. 45 was omitted on 108 of the 25,000 study cards. The wrong punches on these 108 cards were concentrated toward the end of the card, as indicated by Chart III, which shows the distribu-

tion of wrong punches according to the columns in which they occur. If the 168 cards are removed from consideration, the distribution (not shown) of wrong punches on the remaining 24,832 cards, according to the column in which the error occurred, obviously changes. The average number of wrong punches per card is reduced to 1.36.

CHART III

DISTRIBUTION BY COLUMN OF 2,769 WRONG PUNCHES ON 168 CARDS HAVING NO "SIGNATURE" PUNCH IN THE LAST COLUMN



Note. When the signature punch is missing, usually the puncher has somewhere skipped an entry that was to be punched, or (through fault of the machine) has punched two successive entries into one column of the card.

Cards with errors of type iv constitute about 24 per cent of the 3,535 multiple-error cards among the 25,000 studied. None of these involved more than four columns, and most of them involved only two columns. The proportion of cards with this type of error is very nearly the same in each of the five subsamples of 5,000 cards.

Among the 3,535 multiple-error cards there were 389 of type v. On only 6 cards of the 389, three wrong punches were apparently unrelated, and no card contained more than three unrelated wrong punches. During the period that these cards were punched, an average of 3 cards was wrongly punched out of every 200, or 1.5 per cent. On the basis of simple probability, we might therefore expect about 1.5 per cent of the 25,000 study cards to have 2 or more unrelated wrong punches, and about .0225 per cent of them to have 3 or more unrelated wrong punches. Table II shows the agreement between this theory and actual observation.

The distribution of wrongly punched cards according to schedule line number. Chart IV shows the distribution of the wrong cards according to line number of the population schedule. The population schedule

TABLE II

Cards	Theoretical number	Observed number	Difference	Standard error of the difference
Cards with 2 or more unrelated wrong punches	275.9	290	14.0	19.2
Cards with 3 or more unrelated wrong punches	5.6	6	0.4	2.4
Cards with 4 or more unrelated wrong punches	0.1	0	0.1	0.3

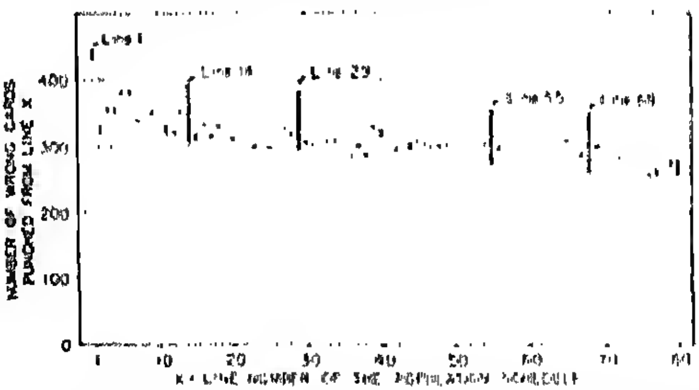
contained 80 lines, i.e., spaces for 80 names, 40 on one side and 40 on the other. Two facts are apparent from Chart IV:

i. There is a gradual decrease in the number of wrong cards for the successive lines of the population schedule.

ii. The number of wrong cards for certain lines is considerably greater than for the others.

CHART IV

THE DISTRIBUTION OF THE 25,000 STUDY CARDS ACCORDING TO LINE NUMBER OF THE POPULATION SCHEDULE



Note.—The ordinate of the upper end of each vertical bar represents the number of wrong cards punched for line x of the Population Schedule. The ordinate of the lower end of each vertical bar represents the number of wrong cards punched for line x, after the 200 cards wrongly punched in col. 44 were excluded. The length of each vertical bar represents the number of cards punched for line x, and having wrong punches in col. 44. The slight downward trend arises from the fact that line x of the schedules was filled more frequently than line $x+1$.

The gradual decrease is to be expected since, on the last schedule that is filled by each enumerator, the early lines are oftener filled than the later ones. Other studies have shown this, and it need not be of concern here. The prominence of the five lines 1, 14, 20, 55, and 68, is, however, more significant. They are sufficiently outstanding to indicate the presence of assignable causes of error. Lines 14, 20, 55, and 68 will be discussed as a group, followed by remarks on line 1.

Lines 14, 29, 55, and 68 are the lines that were designated for supplementary questions on 80 per cent of the schedules.[2] Col. 44 of the punch card is the only column for which the punch is dependent upon the designation of a supplementary line as supplementary. The puncher was instructed to punch "1" in this column for supplementary lines, and to punch "X" for all other lines. Since the lines for which "1" is to be punched constitute only 5 per cent of all lines, it seems likely, on the basis of related coding and punching experience, that the puncher would occasionally by habit punch the usual "X" on a card that required "1."

As a test, then, the 25,000 wrong cards were divided into two groups:

I. Cards that were punched for lines 14, 29, 55, and 68;
II. Cards that were punched for all other lines.

In the first group, one out of every four cards had an error in col. 44, whereas in the second group only one out of fifty had an error in this column. Furthermore, the number of cards with errors in col. 44 for any one of the four lines was at least 4 times the number of cards with errors in col. 44 for any one of the remaining 76 lines of the schedule. Chart IV shows that when cards with errors in col. 44 are excluded, lines 14, 29, 55, and 68 are no longer prominent. We conclude, therefore, that the excess of wrong cards for lines 14, 29, 55, and 68 is due to errors in col. 44.

The errors on cards punched from line 1 of the population schedule do not appear to differ in kind from errors on the other cards. Unlike the group of wrong cards for lines 14, 29, 55, and 68, no particular column seems to account for the errors on the cards for line 1. Now line 1 is the first one that is punched after a break in the puncher's rhythm (removing a completed schedule, and inserting a new one in the holder), and if this is the correct explanation, one would expect to find more than the usual number of errors also for the following types of cards:

i. Cards for line 41. Before punching line 41, the puncher must turn the schedule over in the holder.
ii. Cards for line 1 of the first schedule in a folio. The puncher stops to hand in completed work before commencing a new folio.

With regard to line 41, an excess does occur (see Chart IV), although it is not as much as for line 1. This seems reasonable: the interval is shorter for line 41 than for line 1. With regard to the second point, the interval (time out) preceding line 1 is longer for the first schedule in a

[2] Frederick F. Stephan, W. Edwards Deming, and Morris H. Hansen, "The Sampling Procedure of the 1940 Population Census," this JOURNAL, Vol. 35, December 1940, pp. 615–630.

folio than for any other schedule in the folio. This longer interval was indicated as the cause of the prominence at line 1 when it was found that there were more than twice as many wrong cards from line 1 of schedule 1 as from line 1 of schedule 2 (the schedule that contributed the second largest number of wrong cards).

The frequency of a code in relation to wrong punching. Early in the course of this study, there appeared a principle that seems basic. For some areas, and for some fields of the Population A card, there may occur what we call a "usual" punch. This is a punch which occurs much more frequently than any other punch in the column or field. For example, the "usual" punch in Field 16 (citizenship) is "native," and the other possible punches in this field are relatively rare. When a punch

TABLE III

Errors in Field 10 (Color)

	Color to be punched		
	White	Nonwhite	Total
Cards correctly punched in Field 10	21,029 (98.0%)	3,119 (88.1%)	24,148
Cards incorrectly punched in Field 10	431 (2.0%)	421 (11.9%)	852
Total	21,460	3,540	25,000

Errors in Field 12 (Marital Status)

	Marital Status to be punched		
	Single or married	Any other status	Total
Cards correctly punched in Field 12	20,565 (95.0%)	2,924 (87.5%)	23,489
Cards incorrectly punched in Field 12	1,093 (5.0%)	418 (12.5%)	1,511
Total	21,658	3,342	25,000

Errors in Field 16 (Citizenship)

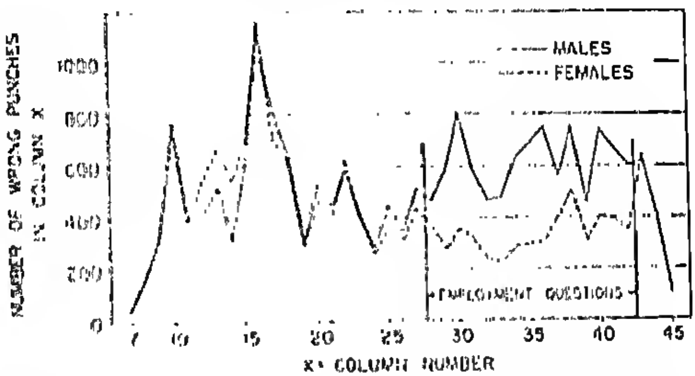
	Citizenship to be punched		
	Native	Any other citizenship	Total
Cards correctly punched in Field 16	22,360 (98.9%)	1,814 (76.2%)	24,174
Cards incorrectly punched in Field 16	260 (1.1%)	566 (23.8%)	826
Total	22,620	2,380	25,000

operator encounters one of the rare entries, she may from habit simply punch it as usual. Thus we expect to find a larger proportion of wrong punches among those cards on which "unusual" entries should be punched. This expectation is borne out by our study of the 25,000 wrongly punched cards.

Table III demonstrates the operation of the principle. In Field 10 (color), "white" may be considered the usual punch. Of the 21,460 cards punched for white persons, only 2.0 per cent were incorrectly punched in the color column; of the 2,540 cards punched for nonwhite persons,

CHART V

THE DISTRIBUTION OF THE WRONG PUNCHES ON THE 14,131 STUDY CARDS PUNCHED FOR MALES AND ON THE 10,869 STUDY CARDS PUNCHED FOR FEMALES



11.0 per cent were incorrectly punched in the color column. Thus it appears to be more than 5 times as easy to make an error in the color column for a nonwhite person as for a white person.

A similar situation holds in the case of Field 16 (citizenship). Among the 25,000 study cards, only 1.1 per cent of those for native Americans showed the wrong citizenship, but 23.8 per cent (or, relatively 20 times as many) of those for persons not native Americans showed the wrong citizenship.

The results of the principle just described are sometimes not so obvious. It will be noted that 14,131 of the 25,000 cards were punched for males. This large proportion is easily accounted for, when it is observed that cards punched for females will often have a "usual" entry in the employment fields (cols. 28–43, inclusive). These "usual" punches are "unknown" and "zero." Males will not so often have these entries. Investigation showed that the larger number of cards for males was a con-

sequence of wrong punches in the employment columns (see Chart V).

In the same way that the action of the principle described above results in an under-representation of females among the wrong cards, it also results in an under-representation of children under 14 years, since they also will have "usual" entries in the employment fields.

In the preceding discussion we have shown that an "unusual" entry is oftener punched incorrectly than a "usual" punch. It was further determined that when an "unusual" entry is punched incorrectly, it is much more likely to be punched as the "usual" punch (if any) than as any other punch.

The actual labor of drawing the sample (the 25,000 study cards), examining them, compiling and studying the figures, was carried out by various staff members, most of it by or under the immediate supervision of Mr. Samuel W. Greenhouse.

CORRELATIONS BETWEEN FUNCTIONS OF VARIABLES

By G. A. Baker
University of California at Davis

The question of the correlation between functions of variables has been raised by Dr. A. E. Brandt.[1] He considers the simple case of the correlation between x_1 and (x_1+x_2) and erroneously comes to the conclusion that such a correlation can never be zero and that the ordinary tests of significance for correlation coefficients cannot be applied. The purpose of this note is to show that the question of correlating a part with a whole is a very restricted case of a very general situation that is of wide practical and theoretical interest and that in many cases the correlations between functions of variables should be computed and compared with zero or their expected value by means of the usual tests of significance.

Sometimes a deliberate search is made for a function of one or more variables that is uncorrelated with another function of one or more variables. As an example, we shall consider the correlation between x_1 and (x_1+x_2) and show that the correlation can be zero in this case. We shall, for the moment, consider only variables that are normally correlated. Consider the populations

$$Z = \frac{1}{2\pi\sigma_1\sqrt{\sigma_2^2-\sigma_1^2}}\exp\left(-\frac{\sigma_2^2}{2(\sigma_2^2-\sigma_1^2)}\left(\frac{x_1^2}{\sigma_1^2}+\frac{x_2^2}{\sigma_2^2}+\frac{2x_1x_2}{\sigma_2^2}\right)\right),\qquad \sigma_2>\sigma_1. \quad (1)$$

Thus x_1 and x_2 are two normally correlated variables with their correlation coefficient equal to $-\sigma_1/\sigma_2$. The populations (1) are not isolated, trivial ones but consist of an infinite class, i.e., of populations for all values of σ_2, σ_1 below the line $\sigma_2=\sigma_1$ in the first quadrant of the (σ_2, σ_1) plane. Now make the transformation of variables

$$u = x_1 \qquad x_1 = u \qquad dx_1dx_2 = du dv$$
$$v = x_1 + x_2 \qquad x_2 = v - u$$

and (1) becomes

$$Z = \frac{1}{(2\pi\sigma_1\sqrt{\sigma_2^2-\sigma_1^2})}\exp\left(-\tfrac{1}{2}\left(\frac{u^2}{\sigma_1^2}+\frac{v^2}{\sigma_2^2-\sigma_1^2}\right)\right),\qquad \sigma_2>\sigma_1. \quad (2)$$

[1] "Should We Correlate a Part and the Whole?" Statistical Service Sheet No. 2, U.S.D.A. Office of Information, Division of Publications, Washington, D. C., October 10, 1941.

The variables u and v are independent. Since u is a "part" of v we have an example of zero correlation between a "part" and a "whole" which consists of a whole class of populations. In sampling from such populations as (1) it is proper to consider the "part"-"whole" correlation and test its divergence from zero by means of the usual significance tests.

In general, consider two variables x_1 and x_2 measured from their means, with standard deviations σ_1 and σ_2 and product correlation, r. Then the expected value of $x_1(x_1+x_2)/\sigma_1\sqrt{\sigma_1^2+\sigma_2^2+2r\sigma_1\sigma_2}$ is $(\sigma_1+\sigma_2 r)/\sqrt{\sigma_1^2+\sigma_2^2+2r\sigma_1\sigma_2}$ which is zero if $r=-\sigma_1/\sigma_2$ and $\sigma_2>\sigma_1$. In all such cases it is proper to test divergence of sample "part"-"whole" correlations from zero. Such a situation might arise in case two characters were negatively correlated but one character was more variable than another.

The correlation of x_1 with (x_1+x_2) has been discussed at some length because it is a simple example of "spurious" correlation and illustrates some of the considerations in connection with the use of "spurious" correlations. "Spurious" correlation was first discussed by Karl Pearson.[2] Yule[3] takes a different view of "spurious" correlation. He says, "It does not follow of necessity that the correlations between indices or ratios are misleading. If the indices are uncorrelated, there will be a similar "spurious" correlation between the absolute measurements $z_1x_3=x_1$ and $z_2x_3=x_2$, and the answer to the question whether the correlation between indices or that between absolute measures is misleading depends on the further question whether the indices or the absolute measures are the quantities directly determined by the causes under investigation."

Many modern investigators have taken the view of Yule. As an example, W. L. Gaines, C. S. Rhode, and J. G. Cash[4] considered the three variables

$x_1 = FCM$ (eight-months partial lactation milk-energy yields in terms of pounds of 4 per cent milk per day)

x_2 = age in years

x_3 = initial live weight in pounds.

In this case $r_{12}\neq 0$, $r_{13}\neq 0$, $r_{23}\neq 0$. The investigators deliberately sought a function of x_1 and x_3 that would not be correlated with x_2. Such a function turned out to be x_1/x_3. To test whether the variable x_1/x_3 was or

[2] "On a Form of Spurious Correlation which may Arise when Indices Are Used in the Measurement of Organs," Proceedings of the Royal Society, Vol. LX, 1897, p. 489.

[3] "An Introduction to the Theory of Statistics," C. Griffin & Co., Ltd., London, 10th edition, revised (1932), p. 216.

[4] "Age, Live Weight and Milk-Energy Yield—A Correction," *Journal of Dairy Science*, Vol. 25, No. 1, 1942, pp. 15-18.

was not correlated with x_2 the deviation of the resulting correlation coefficient from zero was considered.

In general, instead of discussing the correlation between x_1 and (x_1+x_2) we might talk about the correlation between x_1 and some other expression in x_1 and x_2, say $f(x_1, x_2)$. To generalize still further, we might consider the correlation between $g(x_1, x_2, \cdots)$ and $f(x_1, x_2, \cdots)$ where $x_1, x_2, \cdots$ are some measured variables which may or may not be correlated. In the discussion of Yule and Pearson $g(x_1, x_2, \cdots)=x_1/x_3$ and $f(x_1, x_2, \cdots)=x_2/x_3$. In the dairy example $g(x_1, x_2, \cdots)=x_2$ and $f(x_1, x_2, \cdots)=x_1/x_2$. In some cases we may be interested in the correlations between the x_i's and then the correlations between g's and f's might be termed spurious. In other cases we may be interested in the g's and f's or in the particular g's and f's that are not correlated and then the correlations between the x_i's would be termed spurious. To test whether we actually have g's and f's that are uncorrelated we test the significance of the deviation of the observed correlation from zero.

It has been shown that the point of view and purposes of the investigator determine whether or not "part"-"whole" correlations should be computed and the significance of their deviations from zero be tested. To condemn utterly the use of "part"-"whole" correlation is an error that should not be allowed to creep into authoritative statistical writings to add to the confusion of the statistical practitioners. Even though the expected value of a "part"-"whole" correlation is not zero it may still be of value to compute such correlations if the investigator has clearly in mind what he is doing. For instance, in a recent investigation of asparagus at the Davis, California, Experiment Station the "part" yields for 1, 2, $\cdots$, 9 years were correlated with the "whole" 10-year yields for different classes of plants. The expected value of the correlations, if no correlation existed between the yields for different years and the standard deviations were the same for different years, was easily seen to be $\sqrt{x}/\sqrt{10}$, $x=1, 2, \cdots, 9$. The observed values of the "part"-"whole" correlations for the different classes of plants behaved differently with respect to these expected values. This difference in behavior was of great practical and theoretical interest.

A COMMENT ON DEMING'S CLASSIFICATION OF PROBLEMS OF INFERENCE

By Leonard A. Salter, Jr
University of Wisconsin

Mr. Deming's article in the June issue of this Journal contains numerous observations of great importance to research workers, and especially to social scientists. No major exceptions can be taken to the thesis which is set forth. It might be suggested, however, that the implications of the article are much greater even than the paper would lead one to believe. Furthermore, for social scientists at least, the central import of Mr. Deming's statement is that it exposes a basic error in the type of research procedures which have been emphasized during the past twenty years, and it calls for a full reappraisal of customary techniques.

The Type A "problems" to which Mr. Deming refers are actually but problems of *counting*. They are not problems the solution of which is necessary in order to know what to do in order to put action under purposive control. They are questions of counting those elements in a static situation that have previously been determined as strategic, in order to apply an already-decided-upon remedy. The crucial points in the illustrations of Type A social "problems" (the allocation of funds by states or conscription by age groups) are these: that the existential social problem has somehow been defined beforehand, and the basic solution of it has already been decided upon. All that remains for the statistician to do is to count out the elements involved for administrative purposes.

As Mr. Deming says, it is only Type B analyses that can show how a given condition "got that way, or what it ought to be or might have been." Some people would, with excellent reason, insist upon adding that it is just these questions which constitute scientific inquiry. One can only rejoice over Mr. Deming's basic introductory remark that scientific data are collected, not for museums, but to provide recommendations for action. But this is also saying that a science moves forward on Type B achievements. The point about measuring Mr. Deming's table is not merely that in stating its measurement one is predicting that its length will be six feet: the important thing is that a *scientist* measures the table because he is trying to do something with the table. And this means that as a scientist he is really interested in some Type B problem and not merely in the Type A process of measuring the table.

The investigational procedures which Dr. Shewhart classifies as

"classical statistical theory" are the procedures that are directly applicable *to the kind of "problems"* which Mr. Deming classifies as Type A. Now then, when the very limited nature of Type A "problems" is fully grasped, social scientists are found to be in a pretty fix. The resolution of a confused situation into a defined problem and the resolution of a defined problem into a solution that can be used in purposive action—*these are inherently what Deming labels as Type B problems*. Type A questions become a matter of interest primarily when the problem has been defined and the solution already accepted, and this point can be achieved only by Type B investigations.

Mr. Deming is of course quite correct in asserting that Type A procedures can often help to establish partial evidence for a Type B study, but this hardly needs to be underscored today. What rather needs to be emphasized more is that procedures premised on a Type A concept of problems cannot alone give grounded solutions to Type B problems. On this issue, the physical and social scientists stand far apart in practice. The physical scientist can usually operate with his hands and eyes to observe the all-important sequences in his data. He may write up his results so that all his work looks like Type A. But in the laboratory, he *operates* in the manner of Type B. As John Dewey has said, we should pay more attention to what the physical scientist does than to what he says he does.[1] This is the essence of Mr. Deming's illustration about soil treatment experiments.

In the field of social science, the research worker cannot easily *observe operating processes* at first hand. He has to find instances where various social actions have gone through a certain cycle and he has to find the on-going processes of the present. Because these processes may be clearly revealed and exposed at different times and places, because many of them take a long while to unwind, and because many of them are not directly visible to the eye anyhow, the social scientist has to lean heavily on written symbols representing the action in which he is interested. But just because the social scientist makes surveys is no reason for presuming that these are Type A analyses. If they are only Type A analyses, the public needs to be warned of the strict limits on the usefulness of the results. The social scientist uses surveys as often as he does because he has to. Nevertheless, he still has to meet the requirements of Type B analysis by handling his data, however obtained, in a Type B manner. *Thus the question of developing Type B procedures for social studies assumes dominant significance if there is to be any science in social science.*

[1] On most of these questions of scientific inquiry, John Dewey's *Logic, the Science of Inquiry* will be found helpful if not essential.

In my opinion, Mr. Deming's phrase, "repeated patterns in sub-series" reflects the kernel of the procedural methods which need to be developed in social science. The social scientist cannot merely refer to "the underlying forces (in a Type B problem) however derived" even though the statistician may be able to. Nor will it be enough to recommend quick and frequent small samples as long as Type A procedures themselves can only ground solutions of Type A problems, because scientific inquiry has a Type B heart.

It is to be hoped that Mr. Deming's paper will help to force more attention to the basic logical problems of social inquiry and to the development of suggestive aids for the handling of social data in such a way as to preserve the all-important patterns of attributes within the observed cases.[2] If social science is to be able to achieve what Dr. Shewhart has made so much progress in doing for engineering problems, there is plenty of room for those who may care to work on these fundamental problems.

[2] A recent effort in this direction by the writer is "Cross Sectional and Case Grouping Procedures in Research Analysis," *Journal of Farm Economics*, Vol. XXIV, No. 4, November 1942.

CARROLL W. DOTEN, 1871–1942

The death of Professor Carroll W. Doten removed from the roll of the Massachusetts Institute of Technology the name of a distinguished and beloved member of that body. During his long career as a teacher in his chosen fields of economics and statistics, he not only established himself in the hearts and minds of the thousands of students who passed through his classes, but also attained a place of honor among his colleagues on the staff. The wide range of interests which claimed Professor Doten's thought and effort led him into many activities, where he served well his community, his state and his country. For Cambridge, he gave service as president of the Associated Charities, president of the Park Commission and president of the Cambridge Club, among many other similar efforts. For Massachusetts, he rendered service as a member of the bureau of immigration, and in numerous other agencies. For his country, he participated in the work of the United States Shipping Board Emergency Fleet Corporation in 1918 and in the work of the United States Census Bureau. In his professional capacity he contributed to the united efforts of both economists and statisticians. He did research work for the Boston School for Social Workers, was Chief Investigator for the Massachusetts Commission on Compensation for Industrial Accidents, was consulting economist for the National Retail Dry Goods Association and was president of the American Statistical Association.

A native of Vermont, Professor Doten retained throughout his entire life a deep devotion to the state of his birth and to the University of Vermont, from which he received the Bachelor's degree in 1895. During the closing years of his life, following his retirement from Technology, he returned to his native state, to enjoy the evening of his days among the associations of his youth. There in his seventy-first year, he passed away, leaving a great circle of friends, to whom his genial smile and encouraging word had been a joy and inspiration for many years.

FLOYD E. ARMSTRONG

BOOK REVIEWS

GLENN E. MCLAUGHLIN
Review Editor

The Long-run Behavior of Costs in a Chain of Shoe Stores: A Statistical Analysis, by Joel Dean and R. Warren James. Studies in Business Administration, Vol. XII, No. 3. Chicago: The University of Chicago Press. 1942. x, 54 pp. $1.00.

This book attempts to determine statistically the effect of changes in the size of retail shoe stores on their operating costs. This knowledge is sought partly in understanding the theory of long-run cost behavior of an individual firm. Administratively, it is sought for the help it might afford in the formulation of a rational policy of store size and location.

The study is not a long-term analysis as might be expected from the statement of aims but a study of the operating costs of fifty-five (out of several hundred branches of a retail shoe concern) selling men's, boys' and girls' shoes during 1937 and 1938. The authors dismiss the idea of making an investigation of the effect on costs of long-term changes in the level of output partly because such an analysis is "likely" to cover only minor changes in size and "seldom large-scale decreases (changes?)." This seems tantamount to rejecting something for being what it is.

The authors contend that the relation between size and cost of retailing by a group of shops selling the same product and identical in every way except size can be considered as the equivalent of a series of successful adaptations by a retail shop in size and costs to changes in the volume of demand. Perhaps because all of the shops are in one city, the authors fail to deal with the objection that store location might affect the validity of their position although later they show that the daily and weekly sales loads of business-district and residential-district stores in the same city differ considerably. The two authors argue that they need not exclude the business district shops from their analysis as their inclusion only weakens the force of their findings. Apparently, they do not realize that they are dealing with two distinctly different types of retail shop problems whose costs have little relation to each other. They also fail to deal with the objection that a retail chain may often keep poor-paying shops open because of the good will and indirect advertising value they may have. Moreover, the method of allocating central costs such as administrative, insurance advertising and indirect general expenses, rent, and some window display costs tends to hide the true burden of operating the small marginal shops. Where such costs are allocated in proportion to the sales of each shop to the total, some of the economies of bulk purchases and large-scale operation (largely contributed by the bigger shops) are transferred to all of the shops equally.

The most important objection to the authors' method of conducting the statistical analysis is the fact that they chose "a group of stores selling differ-

ent proportions of various types of shoes" whose sales as a total for all shops was "representative of the sales volume of the stores in the chain as a whole." To the reviewer, this would seem to have destroyed their basic assumption that the products of the stores were homogeneous: they have deliberately sought heterogeneity.

The physical volume of sales is chosen as the unit of measurement of shop sales although this measure excludes the sales of rubbers, hosiery, and sundries. In choosing this scale it is recognized that it gives equal weight per pair of shoes sold to the time spent in selling men's shoes and the cheaper shoes for juveniles, although the latter class of sales consumes a larger amount of time per sale. The reviewer feels that if this unit of measurement is adopted, it should exclude boys' and girls' shoe sales and cover men's shoe sales only. However, the reviewer regards the physical volume of shoe sales (in the absence of physical volume figures for rubbers, hosiery, and sundries) weighted by the average time spent in selling each pair as a much better index particularly where, as in this case, heterogeneity in the nature of sales between shops is deliberately sought. It would not have been a difficult matter to have made studies of the average amount of time spent in selling each type and style of shoe and to confine the study to those stores where salesmen spent approximately the same amount of time in making each type of sale.

In measuring cost, two elements of total cost are excluded as tending "to obscure the relations which it is desired to investigate." One of these elements is general administrative expense, the other general indirect expense. The method used in allocating these costs is not stated, but if the volume of sales was the basis then for the sake of consistency, advertising, insurance, rent, and a part of window display charges ought to have been excluded as well.

Correlations are drawn between the physical volume of shoe sales and each element, groups of elements, and total operating costs. The results show that "the elasticity of total costs is increasing over the whole range of output," and is due almost entirely to salesmen's salaries. Some of the groupings of components are open to serious questioning.

In the theoretical section, an attempt is made to explain why the empirical *cost function does not bear out the traditional umbrella-shape explanation* beloved by theorists explaining the relation between costs and output. The function is shown to consist of an ascending series of discontinuous planes. The explanation of the cost line is first considered as temporary departures from an umbrella or parabolic curve and rejected as completely unrelated to any curve series. The shape is explained as the result of discontinuities in costs arising from increases in output beyond the optimum output at that price. Any attempt to determine the shape of the cost line for decreases in output below each optimum level is expressly disclaimed.

MICHAEL DALY

Washington, D. C.

Home Propaganda, Change No. 2. A Report prepared by Mass-Observation for the Advertising Service Guild. London: The Advertising Service Guild. 1942. vi, 78 *pp.* One shilling.

This stimulating booklet presents an analysis of the effectiveness of propaganda directed by the British government toward the home front during the first two years of the present World War. The study was prepared for an association of advertising agents by the staff of Mass-Observation, directed by the energetic and resourceful Tom Harrisson. Although some of the basic facts were secured from the numerous part-time investigators who cooperate with the Mass-Observation organization, most of the primary data were obtained by whole-time workers who made door-to-door calls in London, Worcester, and a group of Kentish villages—324 interviews in each area.

According to this investigation, large governmental expenditures, including over half a million dollars a month for poster and newspaper advertising alone, often had negligible results and occasionally even developed antagonism rather than cooperation. As might be expected from observation of similar activities in the United States, British propaganda campaigns are frequently confusing and even contradictory, as when a single issue of a farm journal contains a plea from one governmental agency for farmers to sell their straw for paper making, along with advice from the ministry of agriculture to pulp the straw for feeding cattle.

The investigation also indicates the dangers of propaganda saturation and the relatively greater returns from appeals "at the patent medicine level" than from those expressed in the dignified language of officialdom. In a democracy at war, it is pointed out, effective propaganda becomes a necessity for persuading people to follow their leaders quickly and willingly.

The report does not, however, over-emphasize the significance of propaganda. "The propaganda of events is more potent that the propaganda of words." A healthy scepticism is also exhibited concerning interviews, since they reveal "what is on the public tongue rather than in the public mind."

On the other hand, the study does reveal a lack of awareness of problems of sampling and reliability. For example, no mention is made of the fact that 324 interviews is a more adequate sample for the "group of Kentish villages" than the same number is for London. Similarly, the evidence that publicity campaigns actually decreased knowledge of governmental programs should be accepted with caution in view of the usual willingness of the Englishman to cooperate, even though grumblingly, with officials. Obviously, Mass-Observation would benefit from having a competent statistician on its staff.

Yet in spite of its possible errors of reliability this study does reveal the futility of much government propaganda with its ill-timed and conflicting appeals. As in other publications by Mass-Observation, the data are presented with clarity and vigor. The report should be carefully studied by social scientists interested in the control or measurement of public opinion

and especially by the officials responsible for the present flood of governmental propaganda for the guidance of the American people.

H. C. BREARLEY

George Peabody College

Parity, Parity, Parity, by John D. Black. Cambridge: The Harvard Committee on Research in the Social Sciences. 1942. xi, 367 pp.

This book treats not only of "parity" between agriculture and non-agriculture, but of price policy in wartime and peacetime as well. The manuscript left the author's hands last May. But the substance remains timely, and will continue so for as long as the parity argument continues to be used in support of legislation aimed to enhance farm prices and farm income. The twenty-four chapters and two forewords are packed with historical facts noted at first hand, replete with ingenious statistical attempts to calculate rational parity standards, full of the shrewd observations on men and events that we have come to expect from Dr. Black. Style and organization—including italics and asterisks—are deliberately journalistic, and to this reviewer more confusing than helpful; but the index helps. Some of the charts are not easy to read on the printed page. The chapters on "Alternative Parity Standards" and "Parity, Please" come closest to summaries of conclusions and warrant the closest reading, although the author describes his conclusions as "relatively the least valuable part of the effort."

Dr. Black is apparently troubled because notions about parity, so useful in gaining support for legislation that he would no doubt call "constructive" in *times of peace, are now obstructing the war effort*. His concern is justified. If anything is needed in furtherance of war effort now, it is sub-parity for some producers—of wheat, for example—and super-parity for others.

The author apparently dreads inflation more than the reviewer does. The general price freeze of April 28 is spoken of as "right in its objectives." The only reasonable price objective to the reviewer is inflation at a rate as slow as can be engineered while stimulating production for war. If painful postwar deflation must follow, it is nevertheless preferable to defeat. Production for war comes first, postwar deflation or not, parity or not.

Obviously the author does not agree with President Roosevelt in regarding parity for farmers based on 1910–14 as an "excellent objective," now or earlier. He suggests other bases as preferable (chapter xii); and so they are if parity must indeed be lived with. He "holds to the judgment that Agriculture during the whole 72 years has been underpaid relative to both Labor and Capital." But he fails to point out that such a situation is inevitable and desirable in the course of a nation's rise from poverty to riches. There must be an incentive to enlarge the proportion of population engaged in non-farm occupations; and where is the incentive to be found unless real incomes are smaller in farming than out?

Obviously also, Dr. Black believes that "the essential principle of parity is sound." In the reviewer's opinion it is not, even in times of peace. The notion that "parity" between "agriculture" and "non-agriculture" ought to be an appropriate object of national economic policy is in fact the second biggest gold brick that has been sold to the American public in many years. Only protectionism is a bigger one—and it is the parent of parity.

In his concluding remarks on "Parity, Please," the author observes, "It is more than possible that within a year an aroused public will wipe the slate clean of all parities . . . ." This is a consummation devoutly to be wished. It seems hardly to be expected, however, unless the public is willing to forsake its current idol of "equal returns for equal effort," returning to the older ideal of remuneration for individuals according to what society will pay for goods and services produced in a competitive system where property rights are respected, or fixing its affections on the current aim of getting the necessary goods produced so as to win the war.

M. K. BENNETT

Food Research Institute
Stanford University, California

The Control of Customer Returns, by Edgar H. Gault and Charles S. Goodman. Ann Arbor: University of Michigan, School of Business, Bureau of Business Research. 1942. 107 pp. $1.50.

Customer returns is an ever present problem in the retail field. Buyers blame salespeople for poor selling; salespeople maintain that customers should make up their minds when purchasing; customers criticize management if merchandise is refused for credit; and management blames buyers, salespeople, and customers when returns become an issue. To pin the tail on the donkey when blindfolded to all the facts is recognized by any retailer as a difficult task.

The Control of Customer Returns is a one-hundred page monograph in which the facts surrounding the return of merchandise are made clear. The study does three things. First, it discusses the nature of the return problem under the familiar headings of procedures in handling returns, causes, and costs. The relation of returns to markdowns and the philosophy of the approval sale add new substance to an old subject.

Second, it indicates the policies for handling returns by calling attention to each store's responsibility in doing a better operating job. These policies, four in number, may or may not be a part of a city-wide plan but should be the basis on which an individual store operates.

Third, it discusses some of the procedures that can be used by stores in controlling customer returns.

Customer returns is not a new subject, but this study has a realistic point of view which should attract more than passing attention. Many retailers and commentators consider customer returns to be a liability. In this report

the liability is acknowledged, but emphasis is placed on "the liberal offering of the return privilege as a powerful promotional device. As a fundamental part of the merchandising operation of a store this policy is expected to increase net profit, while skilled management keeps returns under control through the various techniques which have been described and evaluated." Examples are given to substantiate the argument.

Controlling returns, not attempting to eliminate them because the store shares some of the blame, maintaining that cooperative effort is of questionable value and that each store has its own job to do are important subjects in the report.

Objections to the fact that the study centered in Detroit should not be considered. The work done in other cities has been weighed carefully and the authors' understanding of the subject is too comprehensive to be discounted on this basis.

Here and there a statement is open to question. "The differences observed between the proportion of charge and cash sales returned should not be attributed entirely to the fact of credit purchasing" is controversial. Many purchases are charged which could never be made if they were paid for at time of purchase. Again, "from our own point of view only, the typical customer is not likely to make purchases which are certain to be returned" might be construed as an over-eagerness to prove a point.

None of these criticisms detracts from the real value of the study. Some merchandise men will not agree with the report, but if a retailer who has long considered himself a martyr to customer returns will practice the policies suggested he will have integrated an operation in his business which should have been a part of it long ago.

BISHOP BROWN

Research Bureau for Retail Training
University of Pittsburgh

An Objective Method of Sampling Wheat Fields to Estimate Production and Quality of Wheat, by Arnold J. King, Dale E. McCarty, and Miles McPeek. U. S. Department of Agriculture, Washington, D. C. Technical Bulletin No. 814. 1942. 87 pp. 10 cents.

The bulletin treats of the problem of estimating the production and quality of the wheat crop according to crop-reporting districts of groups 10–20 counties within the wheat-belt States, in the light of the 1939 and 1940 field surveys.

The first section describes techniques employed in field and laboratory. Field work had been carried out by three (occasionally four) crews, each of two men in a car equipped with a crop meter. Assigned to work particular districts each crew selected its own itinerary so as to cover as nearly as practicable a grid-like pattern of highway, roughly equal to the length plus width of the counties individually. Frontage of wheat and total mileage

were recorded, and at intervals of ½ to 4 miles the nearest field was sampled by drawing therefrom two sampling units of 1/2000 acres each. An interesting comment is that the entire area clipped in Kansas in 1940 amounted to only 0.41 acre with an estimated standard error of mean yield per acre of less than 0.4 bushel.

Sample heads were sealed in prepared envelopes together with other information and mailed twice a day to the central laboratory for counting, measuring, threshing, and milling. The plan provided for timely "flash" reports on the production and quality. The average lapse of time in Kansas in 1940 between field sampling and release of information in mimeograph form was 6.4 days.

The second section deals with the efficiency of the sampling and costs for the information obtainable; and sources of bias. Analysis disclosed that the sampling interval and total mileage driven in each district should have been increased, while fewer fields need have been sampled. Another case study in sampling is thus added to the accumulating evidence that the aggregate of small samples, each of one or two sampling units from every subdivision of land area containing the crop, provides the estimate of greatest efficiency.

The authors find that it is more difficult to reduce systematic errors of observation to an acceptable minimum than the random errors of sampling.

The third and final section deals with forecasting the crop from sample measurements of head number and plant length taken 3-4 weeks before harvest. The method is not promising, however, because the regression equations differ from year to year.

The bulletin should prove very helpful in designing sample surveys of other annual crops and of perennial crops such as timber. Proofreading, however, seems not to have been taken very seriously. Rarely indeed does one find typographical errors in bulletins of the U. S. Department of Agriculture; yet this one contains several serious ones, as in the formula on page 18 and 23 and the italicized heading of the fourth column on page 25.

F. X. SCHUMACHER

Duke University

Some Factors Affecting the Supply of Milk and Milk Products in Nova Scotia, by Wilfred J. Garvin. Washington: The Catholic University of America Press. Studies in Economics, No. 4. 1941. viii, 155 pp.

This book "confines its attention to the dairy industry of Nova Scotia, the most populous of the Maritime Provinces . . . The interest of the present study centers around the fact that the total volume of milk production in Nova Scotia is substantially less than the total disappearance of milk and milk products within the Province. Consequently, it is chiefly concerned with the factors conditioning the volume of production, farm costs, and farm income."

Following an introductory chapter on the historical development and

present economic significance of the dairy industry in Nova Scotia, a brief review of some theoretical factors affecting milk supply and a description of the markets for different milk products is presented. Attempts are made to relate, by multiple regression analysis, butter production to various factors that might be expected to affect supply. Location of creameries is indicated and some idea of the extent of overlapping of the supply areas of these creameries is given. Estimates of costs of cream hauling and of plant operation are given.

The data on the costs of cream hauling and plant operation suffer, as do other parts of the book, from a lack of detailed factual information. Only summary figures are given and these appear to represent only the estimates of creamery managers as to their average unit costs, with no analysis of the items included in the total or the methods used to compute them.

On the basis of the analysis presented, it is concluded that "the solution of the present problem of low production, relative to local disappearance, is probably not to be advanced materially by such reductions in cost and changes in creamery organization as are practicably possible. Any economies that may be effected in this field would undoubtedly strengthen the competitive position of Nova Scotia farmers but they would probably be too small to constitute a substantial contribution to the solution of immediate problems." This statement as it stands may be true, but studies completed elsewhere have estimated the potential savings from route reorganization and adjustment of plant capacities to volume to be quite significant and worth securing.

Finally, after attention is paid to factors affecting hay prices and production costs, the following conclusions are reached: (1) that some opportunities exist for reducing the unit cost of creamery operations and for maintaining a grading system for cream purchased; (2) that unit costs of production should be reduced by (a) improving soil fertility, (b) enlarging cultivated acreages, (c) diversifying crops, and (d) utilizing land more intensively.

When the author refers to "the deficiency in the production, relative to provincial disappearance, of milk products in Nova Scotia" as a "problem" some may disagree. Indeed, it would appear that the real problems are those connected with farm organization and operation, and only in the event that changes in these would so modify total production as to change this "deficiency" into a surplus, would the relation of quantities produced to those consumed become a problem. In that case, Nova Scotia producers would be selling on a surplus rather than a deficit market and the present premiums they receive over central Canadian markets would be lost.

The study was presented as a doctoral dissertation. It brings together in one volume material on the dairy industry of Nova Scotia scattered in a number of different places. It presents an account of one phase of the economy of the Maritime Provinces of Canada.

ALAN MACLEOD

New England Research Council

Hospitals and Hospital Patients in New York City, by Neva R. Deardorff and Marta Fraenkel. New York: Welfare Council of New York City. Hospital Discharge Study, Volume One 1942. xvii, 209 pp.

This is the first of a three volume analysis of hospital discharges in New York City during 1933 prepared by the Welfare Council of New York City with the assistance of state and city work relief authorities and the Work Projects Administration. Hospital records are kept primarily for administrative purposes and only those who have attempted to utilize such records in a study of illness can appreciate the almost herculean efforts required to abstract and summarize the records of 576,623 discharges.

After a brief discussion of the background and methods, material is presented concerning the hospitals coöperating in the study, the residence of patients and the amount of inter-borough and inter-health area "migration" for hospital service, the diagnosis of the disease or other condition for which treatment was given, certain demographic data, such as sex, age, color, and religion of the patients, the type of accommodation used and the pay status of the patients, the length of stay in the hospital and condition at time of discharge. These data are presented in 58 text and 12 appendix tables.

It is not possible to appraise fairly the results of this study on the basis of only the first of three contemplated volumes. However, a few general comments can appropriately be made at this time.

Certainly at present and probably for many years to come hospitalized illnesses cannot be accepted as representative of illnesses in the general population. Even in cities with the most complete hospital service probably not more than 10 to 15 per cent of all illnesses and 15 to 20 per cent of all disabling illnesses are hospitalized.

Nevertheless, current statistics of hospital admissions including such items as the number of admissions, the nature of the illness, the type of service obtained, and the financial status of the patients would be very useful in evaluating the adequacy of hospital service in spite of the fact that such statistics would not yield representative information concerning even the more serious illnesses.

The present study is confined to discharged patients. Consequently, it does not reveal the number of persons seeking hospital care during the year, that is the incidence of hospitalization, or the average number of patients in hospitals at a given time, the prevalence of hospitalization. Of course, over a sufficient period of time admissions and discharges must counterbalance but for short periods of time this is not necessarily true especially if hospitals for the chronically ill are included.

Perhaps the principal value of this study will be the experience gained in the attempt to utilize hospital records for other than strictly administrative purposes and the analysis of the possibility of initiating hospital morbidity reports. A discussion of these points to be published in Volume 3 will be eagerly awaited.

HAROLD F. DORN

United States Public Health Service

Income as an Index of the Fiscal Capacity of Michigan Counties, by Marvin A. Bacon. Ann Arbor: University of Michigan Press. Michigan Governmental Studies, No. 8. 1941. vii, 78 pp. 35 cents.

This is a welcome addition to the all-too-short list of studies published to date on measurement of fiscal capacity. The large and growing grants-in-aid for relief and other governmental functions make it increasingly important to determine a sound basis for distributing such aid. And the strong preference for equalizing rather than stimulating aids, as evidenced by recent developments in both state school and relief grants, makes it clear that need and fiscal capacity will be used more and more to determine the distribution of federal and state aids for education and welfare.

Measurements of need and fiscal capacity in actual use are far from perfect. There is neither general agreement as to the best tests nor adequate factual data on distribution of wealth and income either among states or within them. And there is danger that vested interests in a faulty formula will freeze an unsatisfactory distribution system unless progress in measuring fiscal capacity at least keeps pace with growing financial aid.

The author has limited himself to measuring the relative fiscal capacity of the counties of Michigan for the purpose of relief grants. The interest of the study lies, however, in the method, which has much wider application. State income has been broken down into agricultural, other private production, government, and all other. Each of these items has been been apportioned among the counties according to indirect measures of income—the amount of income itself not being available. The method varies for each class. Having estimated county income per capita, the author proceeds to estimate per capita "supernumerary income" by subtracting from per capita income a minimum-of-subsistence income corrected for differences in cost of living in the different counties.

Each step of these estimates has been fully described, and the many limitations of the original data, the method of estimating, and the validity of the checks are discussed. The author makes no exaggerated claims either for the method used or the specific results. In this respect the study is one of the best that the reviewer has seen.

The author concludes that supernumerary income per capita offers the best available index of fiscal capacity. He ranks per capita income excluding income from government second, per capita income from all sources third, assessed valuation per capita fourth, and total assessed valuation (the basis of state relief grants in use in Michigan) fifth. In making this selection the author has apparently overlooked one important factor in determining the fiscal capacity of counties, namely that these local governments must depend on a property tax, the base of which is more closely tied to assessed values than income—total or supernumerary. Fiscal capacity must depend on *taxable* income or wealth.

Mabel Newcomer

Vassar College

The Relation of Cost to Output for a Leather Belt Shop, by Joel Dean. With a Memorandum on *Certain Problems in the Empirical Study of Costs*, by C. Reinold Noyes. Technical Paper 2. New York: National Bureau of Economic Research. 1941. 71 pp. 50 cents

This publication is certainly a combination of extremes. The first 52 pages contain the results of a thorough statistical study of the nature of cost curves in a belt shop and the last 20 pages contain a general critical examination of the theory of costs to see if it is sound enough to warrant an attempt to "test its validity" by quantitative analysis. The reviewer wishes that the two parts had been published separately. Even Noyes admits that his memorandum "is not entirely relevant to this particular study," and is really "an examination of the National Bureau's project and its method of attack, rather than one of this particular study." Noyes closes his memorandum with the statement:

> In these remarks it is far from my purpose to discourage efforts to establish by empirical studies truly scientific generalizations as to the relations of costs to output. . . . On the contrary, I consider this subject of costs to be perhaps the most potentially fruitful field that offers itself to real economic tillers of the soil. My only object is to call attention to the arduous job of clearing, draining, ploughing, harrowing, and sowing that must be done before it is going to be worth while to bring out the reaper.

In spite of this statement most readers are certain to reach the conclusion that Noyes thinks Dean used the reaper without waiting for economic theorists to do the "harrowing!"

The reviewer does not agree with Noyes that Dean should have waited until all of the principal economists agree upon a theory of cost. Noyes starts out by raising such questions as: What is the nature of a cost? Is it a concept? Is it a calculation? Or, is it a kind of event which occurs in economic life? However, the reviewer liked Noyes' statements that "economic theory has reneged" upon the issue of "the destination of costs which are not allocated," "the fact that these questions are treated in a certain way in accounting practice or in the theory of costs is not sufficient to settle the procedure for scientific purposes," and "when reducing operations one plant may retain men according to seniority while another retains the most efficient." The last statement has a bearing upon marginal cost.

Dean studied the effects of nine variables which "the management thought might affect cost." He concluded that he should use "a linear total combined cost function in the multiple regression analysis." The findings for combined cost, direct,[1] and overhead[2] cost are presented in both graphic and tabular form. The total, average, and marginal aspects of cost behavior are shown for the three independent variables which were significant. Limited space prevents detailed examination of the results and of the adjustments made

[1] Cement, direct labor, and leather.

[2] Fixed charges, indirect labor, repairs, and supplies.

for changes in wage rates and material prices. Original data were not presented in tabular form, so the accuracy of the work cannot be evaluated.[3] The study appears to be an excellent one. The reviewer would like to read Dean's statement of what changes (if any) he would make in his study if he could repeat it after reading Noyes' memorandum.

HARRY PELLE HARTKEMEIER

University of Missouri

Operating Results of Department and Specialty Stores in 1941, by Charles A. Bliss, and *Department Store Financial Requirements 1920–1940*, by Pearson Hunt and Albert R. Koch, Bulletin No. 114. Boston: Harvard University Graduate School of Business Administration, Bureau of Business Research, 1942. vi, 62 pp. $2.50.

This year's edition of this series is one of the more important, because it serves as a bench mark by which to measure the influence on department store retailing of the General Maximum Price Regulation, of shortages of consumer merchandise, and of other war conditions. This is the twenty-second year the Harvard Business School has published this analysis—which from the start has been a cooperative venture with the National Retail Dry Goods Association. This series, then, offers one of the most authentic and most historically complete series of retailing data available. The cooperation of retail firms—figures are based on reports from 492 companies operating 651 stores, and representing 37.5 per cent of the total volume of sales of department stores and departmentized specialty stores in the United States—and the year-by-year refinements in method of analysis make comments on the adequacy and on the method of treatment superfluous.

A note of caution in the text of the report reminds readers that the data in the report have been submitted, typically, by larger stores, and by stores whose executives consider store operation from the point of view of control. Although the data may not be typical of department store retailing as a whole, it is the feeling of the reviewer that the report reflects operating conditions of the most significant and most influential stores.

It is difficult to summarize summary tables. Several points of general interest should be stressed, however. In 1941, department store and specialty store gross margin was substantially higher than in any of the other twelve comparative years covered in the report; and for department stores, total expense was lower than for any year since 1930, and net profit was substantially higher than for any of the other years shown. However, a significant and sinister increase in the total expense per transaction appeared; only increased size of transaction and increased numbers of transactions made the profit showing possible. If department stores suffer a decrease in the average amount of sale, or in the number of sales, it seems likely that they will lose their recent gains.

[3] In Table 9 (p. 34) 0.28 should be changed to .028. If the material on page 41 does not contain a misprint it needs clarification.

Financial Requirements of Department Stores, 1920–1940 (added to the report) is a cooperative study with the National Bureau of Economic Research. This report analyzes the financial statements of fourteen large department stores over a twenty year period to ascertain the sources of funds for store operation. Figures used for this analysis are appended to the report. Conclusions are that the growth of consumer instalment credit has pushed a growing financing burden on department stores, that inventories are being held increasingly to close relationship with sales, that an increasing proportion of reserves is being held in cash, but that the proportion of working capital to total assets has decreased; and that fixed assets have increased over the twenty-year period both absolutely and in relation to total assets.

LAWRENCE C. LOCKLEY

The Curtis Publishing Company

French Predecessors of Malthus, by Joseph J. Spengler. Durham: Duke University Press, 1942. ix, 398 pp. $4.50.

As the author states in the preface, there has been a need for a detailed and authoritative study of French population and wage theory of the pre-nineteenth century period, especially since many of the original materials are not readily available and the secondary sources are generally too specialized or too summary in character. This need is well filled by the present volume which gives detailed summaries of both major and minor French contributions to early population literature, being in effect a fuller documentation of the author's earlier work, *France Faces Depopulation*.

The scope of coverage is greater than indicated by first glance at the title, for population and wage theory are broadly defined and the writers dealt with are the predecessors of Malthus in point of time, not his precursors in point of thought. To judge from its form of organization the book is designed for reference use, the treatment of the subject matter being neither strictly chronological nor in terms of theoretical sequences. The various writers are dealt with in turn, full abstracts of their works being given, but further organization is secured through their grouping into several schools of thought, numerous cross-references, and a detailed index.

It is shown that although the subject of population received considerable attention from French writers of the seventeenth and eighteenth centuries, few anticipated in any degree the Malthusian doctrine. Opinion in the earlier period is seen to have been quite unanimously in favor of population growth as the sure means to national strength and prosperity. During the eighteenth century the Physiocrats and other economists came to emphasize the importance of resources and industry as sources of wealth, at the same time that the growing liberal and social point of view stressed considerations of individual welfare. Elements of the Malthusian doctrine are shown to have appeared here and there but only a few writers, notably Bruckner and Moheau, were in a real sense precursors of Malthus. Although foreigners such as

Arthur Young could conclude that France was suffering from overpopulation, the existing social ills were currently attributed to faulty government, defective economic organization, and maldistribution of wealth.

A unified theory of wages was not to be developed until later. In the discussion of the subject in eighteenth century France, quite as in contemporary England, there are shown to have been elements of supply and demand or subsistence, wage-fund, and productivity theories.

E. P. HUTCHINSON

Washington, D. C.

On the Principles of Statistical Inference, by Abraham Wald. Notre Dame, Indiana: Notre Dame Mathematical Lectures, No. 1, 1942. 47 pp.

Two problems of statistical inference have received considerable attention during the last twenty or thirty years: the testing of statistical hypotheses and the problem of estimation. In this booklet, containing four lectures delivered by Dr. Abraham Wald at Notre Dame University in February, 1941, these two problems are stated, as well as the general problem of statistical inference of which they are special cases; a brief general review of the solutions obtained for them is given; and an approach to the general problem, following concepts developed by Dr. Wald, is indicated. For the mathematician or statistician who wishes to get a brief over-all review of the progress that has been made in mathematical statistics, and of some of the major problems still unsolved, Dr. Wald's paper may be highly recommended.

The contents are divided into six chapters: the first is an introduction which defines the terminology used, presents the general problem of inference, and gives the special cases of estimation and of testing a statistical hypothesis. The second chapter considers the concepts introduced by J. Neyman and E. S. Pearson in developing their theory of testing an hypothesis regarding the value of a single unknown parameter, and the third chapter gives R. A. Fisher's approach to unique or point estimation of a parameter. Estimation of the value of a parameter by a confidence interval, following Dr. Neyman's theory, is treated in Chapter IV. Unfortunately, the briefness of the publication has led to the omission of many interesting topics, notably testing composite hypotheses, estimating one unknown parameter when there are several unknown parameters involved in the probability distribution, and Fisher's theory of fiducial limits. However, despite their briefness and the consequent omissions, these four chapters give a very good outline of the basic ideas used in the theories of Fisher, Neyman and Pearson. Dr. Wald's own recent work in developing asymptotically most powerful tests and asymptotically shortest confidence intervals is treated in Chapter V, and several interesting results are given. In the last chapter, the author considers the general problem of statistical inference, outlines the steps that must be taken to solve it, and gives his approach to the solution, through the consideration of the risk of loss involved in making an incorrect inference.

Although briefness has led to omissions of subjects and of detailed develop-

ments, this very briefness has the advantage of letting the reader obtain quickly an excellent grasp of the entire field of statistical inference, and of the concepts used in modern statistical theory.

MARK EUDEY

Statistical Laboratory,
University of California,
Berkeley

Changes in Distribution of Manufacturing Wage Earners, 1899-1939. Washington: Bureau of the Census and Bureau of Agricultural Economics, 1942, 268 pp. 30¢.

This report, prepared by Harold D. Kube and Ralph H. Danhof, is a study of the geographic distribution of manufacturing wage earners in the United States and results from the cooperative study that the Bureau of the Census and the Bureau of Agricultural Economics started early in 1940.

The purpose of the study was to develop accurate manufacturing wage-earner data to form the basis for further analytical work. No part of the report is designed to explain the reasons for changes in the distribution of manufacturing wage earners or the economic consequences of such changes. The study is based primarily upon the Censuses of Manufactures and the maps, charts, and tables forming the body of the report are presented in three sections.

Section I describes changes in the distribution of manufacturing wage earners between 1899 and 1939 for the Nation as a whole, the 9 geographic divisions, the 48 states, the 33 industrial areas, and various city groups. The importance of the changes in the distribution of wage earners is indicated by relating them wherever possible with corresponding changes in the population of each area. Sufficient data are found in this section to prepare analytical studies which would trace the growth process of the present manufacturing wage-earner employment in each area.

Section II contains maps, charts, and tables showing changes in the distribution of wage earners between 1929 and 1939 in geographic divisions, industrial areas, and city groups for 24 manufacturing industries considered important to national defense. Wage-earner data are indicated on both the 1929 and the 1939 maps by counties and the relative importance of each county is designated by shading or ruling for 5 class groups. The limits of the categories were set wide enough that information could be given for all counties in which the industry occurs without violating the disclosure rule which forbids the Bureau of the Census to reveal data of a given establishment. It is possible, however, to discern readily whether the industry is of minor, small, medium, considerable, or great importance in terms of number of wage earners in any part of the country. Charts and tables are presented to support the maps by showing the distribution of wage earners by geographic divisions, industrial areas, and 5 groupings of cities for each manufacturing census between 1929 and 1939 inclusive. Such information should contribute much to any regional study of the selected industries.

Section III contains source material in two sets of tables. The first set shows the number of manufacturing wage earners employed between 1899 and 1939 for the United States, geographic divisions, states, counties, industrial areas, and groups of cities. The second set comprises lists of the cities included in the various city groups based on the 1930 and 1900 census of population. The source data on the first set have been adjusted so that they are as nearly as possible comparable over the 40-year period. The authors hope that these statistics will be the basis for future historical analyses of the development of the geographic structure of American manufacturing activity.

The county and city wage-earner data published in this report should be of particular usefulness to statisticians, economists, and other students of manufacturing data because they are probably the only set of data available in which previously published census data were revised to make the county and city statistics more comparable over the 1899–1939 period. Many changes were made in the earlier published data and, in addition, many new figures were added by estimating the number of wage earners for those counties and cities for which separate information was not previously reported.

Changes in scope of the census, definition of average number of wage earners, areas included in city or county boundaries, and coverage of the several censuses were considered in correcting the wage-earner data for comparability. The estimates of county and city statistics to overcome the lack of completeness in published census data were made according to accepted statistical procedures. Good use seemed to have been made of the information available and it is unlikely that the error in these estimates will affect the usefulness of the data. The authors indicate that the possible margin of error is usually rather small because the upper and lower limits of their estimates could be determined with reasonable accuracy.

The report is highly recommended as a source book of statistics to all persons interested in analyzing changes in the distribution of manufacturing wage earners or in making regional studies of selected manufacturing industries. Its small cost for the wealth of available information makes the bulletin especially attractive.

Oscar L. Endler

National Resources Planning Board

Industrial Statistics, by H. A. Freeman. New York: John Wylie & Sons, Inc., 1942. ix, 178 pp. $2.50.

Production could be increased by thirty per cent, and costs could be decreased by a like amount, with a corresponding improvement in quality, if the principles of quality control were to be applied to production, in place of the conventionally accepted inspection methods now used. By quality control is implied establishing a qualis level and so regulating and controlling the tools of production and the manufacturing process so that this qualis is always obtained.

At present the accepted procedure in most industries is to make the product and then submit it to an inspector for acceptance or rejection. This guarantees the consumer a quality product but does not insure the producer against loss from spoilage and defective workmanship. The cost of this spoilage is ultimately absorbed by the consumer which results in higher prices. Production is also restricted by the amount of spoilage since this requires reworking of the rejected product or scrapping of the defective units and reordering new materials in order to make the units over again. This waste of man power and machine hours is irreplaceable.

It has long been recognized that conventional inspection methods, while satisfactory to the consumer, did not contribute to increased production or lower costs for the producer. The amazing thing is that manufacturers have been satisfied to organize and hire a production force and then parallel this organization with an inspection department with authority to reject the product made by the manufacturing organization. It is surprising that the principles of *preventive* inspection have not received more attention in view of the tremendous losses due to spoilage during manufacturing.

Dr. Shewhart, of the Bell Laboratories, recognized this problem a few years ago, and as a result of his pioneering efforts, a considerable amount of work has been done in the last half dozen years to replace the post mortem inspection operation with proper controls during manufacturing so that the desired quality level might be maintained during the manufacturing cycle, thus reducing rejections to a minimum.

In his new book, Professor Freeman has attacked the problem of quality control in a novel but effective manner. He starts out by determining the answer to the question: Of two or more lots of materials, did they all come from the same universe, or, in the language of the shop man, were they all produced under identical conditions? In other words, Professor Freeman gives a mathematical solution for the age-old problem of probability: Is the difference between lots what might be expected due to chance, or are the lots significantly different from each other?

From this test, Professor Freeman proceeds to the use of correlation and the analysis of variance to be applied to manufacturing operations during the manufacturing cycle. Unfortunately, on this subject, Professor Freeman does not have too much to say about the control of quality applied to the metal working industries. However, the quality control engineer will find this study a practical guide, though somewhat too mathematical for the ordinary shop man. Professor Freeman attempts to get around this difficulty by segregating into separate sections the more theoretical treatment of the mathematics of the problem.

The tables at the end of the book are a boon to the serious student of this subject, since in the past he has had to go to have a dozen books to get this information.

Joseph Manuele

Westinghouse Electric and Manufacturing Company

PUBLICATIONS RECEIVED*

Anthony, Donald, et al. *How Collective Bargaining Works.* New York: The Twentieth Century Fund. 1942. xviii, 986 pp. $4.00.

Bagley, William Chandler, Jr. *Soil Exhaustion and the Civil War.* Washington: American Council on Public Affairs. 1942. xi, 101 pp. $2.00.

Beebe, Gilbert Wheeler. *Contraception and Fertility in the Southern Appalachians.* Baltimore: The Williams & Wilkins Company. 1942. xiii, 274 pp. $2.50.

Brown, Earl, and Leighton, George R. *The Negro and the War.* New York: Public Affairs Committee. 1942. 32 pp. 10 cents.

Crum, William Leonard, Fennelly, John F., and Lawrence, Howard S. *Fiscal Planning for Total War.* New York: National Bureau of Economic Research. 1942.

Deardorff, Neva R., and Fraenkel, Marta. *Hospitals and Hospital Patients in New York City.* New York: Welfare Council of New York City. Hospital Discharge Study, Volume One. 1942. xvii, 209 pp.

East, W. Norwood. *The Adolescent Criminal.* London: J. & A. Churchill, Ltd. 1942. xi, 327 pp. 18 shillings.

Faville, David E., Jones, Dix M., and Soule, Richard H. *Merchandise Availability in Utah.* Stanford University: Stanford University Press. 1942. xvi, 229 pp. $3.00.

Field, Kenneth, Editor. *American Finance.* New York: American Finance Association. Abstracts of Papers Presented at the 1941 Annual Meeting of the Association. 1941. v, 67 pp.

Holmes, Maurice C. *Outline of Probability and Its Uses.* Minneapolis: Burgess Publishing Company. 1936. viii, 119 pp. $1.50.

Hoover, Edgar M., and Dean, Joel, Editors. *Readings in the Social Control of Industry.* Philadelphia: The Blakiston Company. 1942. viii, 404 pp. $2.75.

Jacoby, Neil H., and Saulnier, Raymond J. *Term Lending to Business.* New York: National Bureau of Economic Research. 1942. xvii, 163 pp. $2.00.

Mathe, Louis. *Will Electric Power Be a Bottleneck?* Washington: The Brookings Institution. 1942. 67 pp. 50 cents.

Molina, E. C. *Poisson's Exponential Binomial Limit.* New York: D. Van Nostrand Company, Inc. 1942. 47 pp. $2.75.

Palmer, Gladys L., assisted by Cohn, Samuel M. *War Labor Supply Problems of Philadelphia and Its Environs.* Philadelphia: University of Pennsylvania, Industrial Research Department of Wharton School of Finance and Commerce. Research Report No. 5. 1942. vii, 54 pp. 50 cents.

Pollock, James K., and Eldersveld, Samuel J. *Michigan Politics in Transition, An Areal Study of Voting Trends in the Last Decade.* Ann Arbor: University of Michigan Press. 1942. 71 pp.

Stigler, George J. *The Theory of Competitive Price.* New York: The Macmillan Company. 1942. vii, 197 pp. $3.00.

Toner, Jerome L. *The Closed Shop.* Washington: American Council on Public Affairs. 1942. viii, 205 pp. $3.25.

Treloar, Alan E. *Random Sampling Distributions.* Minneapolis: Burgess

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. [illegible] have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications [illegible] *Revue de l'Institut International de Statistique*, *Journal of the Royal Statistical Society*, *Review of Economic Statistics*, *Population Index*, *Transactions of the Actuarial Society of America*, [illegible] *Institute of Actuaries*, and *Sankhyā, The Indian Journal of Statistics*.—Editor.

Publishing Company. 1942. 84 pp. $2.25.

Yoder, Dale. *Personnel Management and Industrial Relations.* New York: Prentice-Hall, Inc. 1942. xxv, 848 pp. $5.65.

Government Publications

U. S. Department of Commerce, Bureau of the Census, and Department of Agriculture, Bureau of Agricultural Economics. *Changes in Distribution of Manufacturing Wage Earners, 1899-1939,* prepared by Harold D. Kube and Ralph H. Danhof. 1942. xv, 256 pp. 30 cents.

Minnesota Resources Commission. *Minnesota Resources, 1938-40: A Report of the Minnesota Resources Commission.* Saint Paul, Minnesota: Resources Commission. Vol. I. 1942. 177 pp.

—. Vol. II. 1942. 209 pp.

—. Vol. III. 1942. 329 pp.

—. Vol. IV. 1942. 54 pp.

Social Security Board, Federal Security Agency, Bureau of Research and Statistics. *Statistics of Family Composition in Selected Areas of the United States, 1934-36. Volume I: Los Angeles, California.* 1942. xxvi, 303 pp.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 37: 1942

NUMBERS 217-220

Published Quarterly by the

AMERICAN STATISTICAL ASSOCIATION

WASHINGTON, D. C.

1942

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

CONTENTS OF VOLUME 37

ARTICLES, NOTES AND DISCUSSIONS

BOOK REVIEWS